

Hippos, nutrient kingpins  
of African rivers p. 802

Engineered DNA as a cellular  
memory device pp. 813 & 825

How mammalian mating  
affects infanticide p. 841

# Science

\$10  
14 NOVEMBER 2014  
sciencemag.org

AAAS

## *Lightning in a warming world*

Projecting more strikes in the  
continental United States p. 851







# Webinar

## Untangling the tumor microenvironment

Illuminating the complex interactions and functions of immune cells

**Wednesday, December 10, 2014**

12 noon Eastern, 9 a.m. Pacific, 5 p.m. UK, 6 p.m. Central Europe

**During this webinar viewers will be able to ask questions of the panelists and will learn about:**

- Cutting-edge research targeting the B-cell receptor signaling pathway that has recently demonstrated therapeutic promise
- Methods that can provide in depth information on cancer phenotypes, including simultaneous immunohistochemistry of multiple biomarkers, multiplexed imaging, single cell quantitative analysis, and automated phenotyping
- How host-tumor interaction analysis in breast cancers could form the basis for assays to guide therapy and monitor response.

Certain antibody therapies have demonstrated the potential for directing a patient's own immune system against tumors. Further advances in this area will depend upon a detailed understanding of the tumor microenvironment and characterization of the location and status of immune cells and their interaction with tumor cells. This will require methods that provide phenotyping of immune and cancer cells combined with information about their spatial relationship in tumor regions. Additionally, a deeper understanding of the signaling cascades active in immune recognition of cancers is crucial. During this webinar, we will discuss the bringing together of multiplexed fluorescent immunohistochemistry, advanced microscopy techniques, and bioinformatics, and how these are now enabling new insights into cancer biology and immunology.

### Speakers



**Scott J. Rodig, M.D., Ph.D.**  
Dana-Farber Cancer Institute  
Boston, MA



**Edward C. Stack, Ph.D.**  
PerkinElmer  
Hopkinton, MA

**REGISTER NOW!** [webinar.sciencemag.org](http://webinar.sciencemag.org)

Webinar sponsored by



Brought to you by the  
Science/AAAS Custom  
Publishing Office



 @SciMagWebinars

# CONTENTS

14 NOVEMBER 2014 • VOLUME 346 • ISSUE 6211



806

Deadly effects of  
neonicotinoid insecticides

## NEWS

### IN BRIEF

**792** Roundup of the week's news

### IN DEPTH

#### **794 RELIEF GREETES ACQUITTALS IN ITALY EARTHQUAKE TRIAL**

But some still fault scientists for hasty risk assessment *E. Carlidge*

#### **795 JUST ONE POLIOVIRUS LEFT TO GO**

Vaccinators may have wiped out the second of three strains of the virus  
*By L. Roberts*

#### **796 FOR SCIENCE, IMPACT OF REPUBLICAN WIN MAY BE MUTED**

Congress could still act on some research-related issues despite worsening relations with the White House  
*By J. Mervis and D. Malakoff*

#### **797 Will pork be back on the menu?**

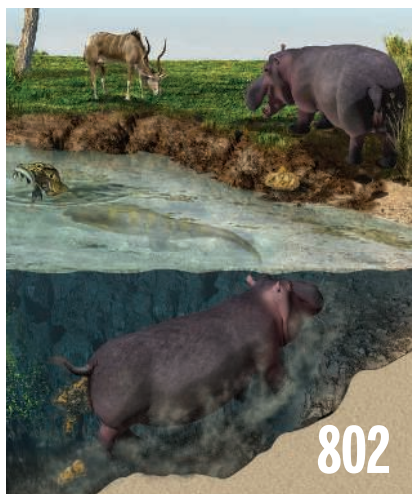
*By J. Mervis*  
► EDITORIAL P. 791

#### **798 NO SEXISM IN SCIENCE? NOT SO FAST, CRITICS SAY**

Study of hiring in math-heavy fields provokes outcry  
*By R. Bernstein*

#### **799 THE GENES THAT TURNED WILDCATS INTO KITTY CATS**

Findings help show how all animals became tame *By D. Grimm*



802

#### **800 THE THEORETICAL PHYSICIST BEHIND *INTERSTELLAR***

What Kip Thorne loved about the new science-fiction movie—and what made him cringe *By D. Clery*

#### **801 CONTAMINATION PLAGUES SOME MICROBIOME STUDIES**

Extraneous DNA can skew results of surveys of low-density microbial communities *By E. Pennisi*

### FEATURE

#### **802 THE RIVER MASTERS**

Hippos are the nutrient kingpins of Africa's waterways  
*By E. Pennisi*

► PODCAST

## INSIGHTS

### PERSPECTIVES

#### **806 THE TROUBLE WITH NEONICOTINOIDS**

Chronic exposure to widely used insecticides kills bees and many other invertebrates *By F. Sánchez-Bayo*

#### **808 HIDDEN FEATURES OF HUMAN HOTSPOTS**

Multiple flavors of genetic changes occur at meiotic recombination sites  
*By B. de Massy*

► RESEARCH ARTICLE P. 826

#### **809 IMPROVING PLANET-FINDING SPECTROMETERS**

Adaptive optics will enhance the precision of exoplanet surveys  
*By J. R. Crepp*

#### **810 MESENCHYMAL PROGENITOR PANOPLY**

Distinct pools of progenitor cells develop stromal tissues of the lung  
*By J.-H. Lee and C. F. Kim*

► RESEARCH ARTICLE P. 827

#### **812 DID ANTARCTICA INITIATE THE ICE AGE CYCLES?**

Gradual buildup of Antarctic ice sheets preceded extensive Northern Hemisphere glaciation about 3 million years ago *By R. McKay*

► REPORT P. 847

#### **813 DYNAMIC GENOME ENGINEERING IN LIVING CELLS**

Engineered gene expression systems provide cells with a molecular memory of their past

*By S. Ausländer and M. Fussenegger*

► RESEARCH ARTICLE P. 825

#### **815 AN ECONOMIC PERSPECTIVE ON THE EPA'S CLEAN POWER PLAN**

Cross-state coordination key to cost-effective CO<sub>2</sub> reductions  
*By M. Fowlie et al.*

### BOOKS ET AL.

#### **817 THE MALARIA PROJECT**

*By K. M. Masterson,*  
*reviewed by I. W. Sherman*



818

#### **818 PRIVACY, BIG DATA, AND THE PUBLIC GOOD**

*J. Lane et al., Eds.; reviewed by F. H. Cate*

### LETTERS

#### **819 FAUNA IN DECLINE: A BIG LEAP TO SLAVERY**

*By Y. J. Masuda and T. Scharke*

#### **819 FAUNA IN DECLINE: MANAGEMENT RISKS**

*By M. D. Smith*

#### **819 RESPONSE**

*By L. Withey et al.*

#### **820 FAUNA IN DECLINE: BEYOND EXTINCTION**

*By J. C. Briggs*

#### **821 FAUNA IN DECLINE: THE COMMUNITY WAY**

*By S. Boyer*

#### **821 FAUNA IN DECLINE: PROTECT FORESTS NOW**

*By J. V. Wells*

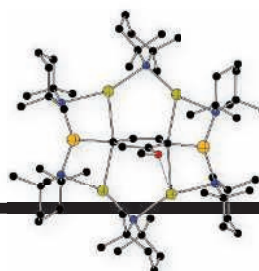
#### **821 ERRATA**





# 808 & 826

Mapping genetic crossover  
in humans



# 834

A base that brings along  
a template

## RESEARCH

### IN BRIEF

**822** From *Science* and other journals

### RESEARCH ARTICLES

#### 825 SYNTHETIC BIOLOGY

Genomically encoded analog memory  
with precise in vivo DNA writing in  
living cell populations

*F. Farzadfar and T. K. Lu*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:

[dx.doi.org/10.1126/science.1256272](http://dx.doi.org/10.1126/science.1256272)

► PERSPECTIVE P. 813

#### 826 DNA RECOMBINATION

Recombination initiation maps of  
individual human genomes

*F. Pratto et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:

[dx.doi.org/10.1126/science.1256442](http://dx.doi.org/10.1126/science.1256442)

► PERSPECTIVE P. 808

#### 827 MESENCHYMAL CELLS

Defining a mesenchymal progenitor  
niche at single-cell resolution

*M. E. Kumar et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:

[dx.doi.org/10.1126/science.1258810](http://dx.doi.org/10.1126/science.1258810)

► PERSPECTIVE P. 810

### REPORTS

#### 828 NANOPHOTONICS

Plasmoelectric potentials in metal  
nanostructures

*M. T. Sheldon et al.*

#### 831 INTERFACIAL WATER

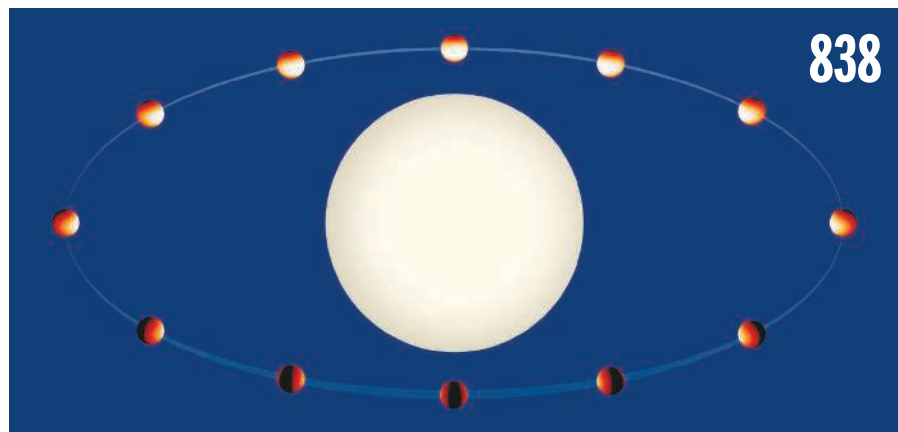
The structure of interfacial water  
on gold electrodes studied by x-ray  
absorption spectroscopy

*J.-J. Velasco-Velez et al.*

#### 834 STRONG BASES

Directed ortho-meta'- and meta-  
meta'-dimetalations: A template base  
approach to deprotonation

*A. J. Martínez-Martínez et al.*



# 838

#### 838 EXOPLANET ATMOSPHERE

Thermal structure of an exoplanet  
atmosphere from phase-resolved  
emission spectroscopy

*K. B. Stevenson et al.*

#### 841 SEXUAL CONFLICT

The evolution of infanticide by males  
in mammalian societies

*D. Lukas and E. Huchard*

#### 844 LANDSCAPE HYDROLOGY

The hydrological legacy of deforestation  
on global wetlands

*C. Woodward et al.*

#### 847 PALEOCEANOGRAPHY

Antarctic role in Northern Hemisphere  
glaciation S. C. Woodward et al.

► PERSPECTIVE P. 812

#### 851 CLIMATE CHANGE

Projected increase in lightning strikes in  
the United States due to global warming

*D. M. Romps et al.*

#### 854 CIRCADIAN RHYTHM

Dysrhythmia in the suprachiasmatic  
nucleus inhibits memory processing

*F. Fernandez et al.*

#### 857 CILIA AND FLAGELLA

A molecular ruler determines the repeat  
length in eukaryotic cilia and flagella

*T. Oda et al.*

#### 861 VIRAL INFECTION

Prevention and cure of rotavirus  
infection via TLR5/NLRC4-mediated  
production of IL-22 and IL-18

*B. Zhang et al.*

### DEPARTMENTS

#### 791 EDITORIAL

Out of sight, out of mind

*By Mary Woolley and Alan I. Leshner*

► NEWS STORY P. 796

#### 882 WORKING LIFE

The point of it all

*By Joseph Swift*

### ON THE COVER



Lightning strike  
above the city of San  
Francisco, photographed  
from the hills above the  
University of California,  
Berkeley, campus. With  
rapid global warming  
projected for the 21st  
century, the frequency

of lightning is poised to increase. For more  
on the future of lightning, see page 851.

Photo: © Douglas Keister/Corbis

|                       |     |
|-----------------------|-----|
| Science Staff .....   | 790 |
| New Products .....    | 868 |
| Science Careers ..... | 869 |

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2014 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscription (51 issues): \$153 (\$74 allocated to subscription). Domestic institutional subscription (51 issues): \$1282; Foreign postage extra: Mexico, Caribbean (surface mail) \$55; other countries (air assist delivery) \$85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request. GST #1254 88122. Publications Mail Agreement Number 1069624. Printed in the U.S.A.

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. Postmaster: Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090-6178. Single-copy sales: \$10.00 current issue, \$15.00 back issue prepaid includes surface postage; bulk rates on request. Authorization to photocopy material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that \$30.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for Science is 0036-8075. Science is indexed in the Reader's Guide to Periodical Literature and in several specialized indexes.



## Editor-in-Chief Marcia McNutt

Executive Editor Monica M. Bradford News Editor Tim Appenzeller

Managing Editor, Research Journals Katrina L. Kelner

Deputy Editors Barbara R. Jasny, Andrew M. Sugden(UK), Valda J. Vinson, Jake S. Yeston

## Research and Insights

SR. EDITORS Caroline Ash(UK), Gilbert J. Chin, Lisa D. Chong, Maria Cruz(UK), Julia Fahrenkamp-Uppenbrink(UK), Pamela J. Hines, Stella M. Hurlley(UK), Paula A. Kiberstis, Marc S. Lavine(Canada), Kristen L. Mueller, Ian S. Osborne(UK), Beverly A. Purnell, L. Bryan Ray, Guy Riddihough, H. Jesse Smith, Jelena Stajic, Peter Stern(UK), Phillip D. Szurmi, Brad Wible, Nicholas S. Wigginton, Laura M. Zahn ASSOCIATE EDITORS Brent Grocholski, Melissa R. McCartney, Margaret M. Moerchen, Sacha Vignier ASSOCIATE BOOK REVIEW EDITOR Valerie B. Thompson ASSOCIATE LETTERS EDITOR Jennifer Sills CHIEF CONTENT PRODUCTION EDITOR Cara Tate SR. CONTENT PRODUCTION EDITORS Harry Jach, Trista Wagoner CONTENT PRODUCTION EDITORS Jeffrey E. Cook, Chris Filiatreau, Cynthia Howe, Lauren Kmec, Barbara P. Ordway SR. EDITORIAL COORDINATORS Carolyn Kyle, Beverly Shields EDITORIAL COORDINATORS Ramatoulaye Diop, Joi S. Granger, Lisa Johnson, Anita Wynn PUBLICATIONS ASSISTANTS Aneera Dobbins, Jeffrey Hearn, Dona Mathieu, Le-Toya Mayne Flood, Shannon McMahon, Scott Miller, Jerry Richardson, Rachel Roberts(UK), Alice Whaley(UK), Brian White EXECUTIVE ASSISTANT Anna Bashkirova ADMINISTRATIVE SUPPORT Janet Clements(UK), Michael Crabtree(UK, Intern), Lizanne Newton(UK), Maryrose Madrid, John Wood(UK)

## News

NEWS MANAGING EDITOR John Travis INTERNATIONAL EDITOR Richard Stone DEPUTY NEWS EDITORS Daniel Clery(UK), Robert Coontz, Elizabeth Culotta, David Grimm, David Malakoff, Leslie Roberts CONTRIBUTING EDITORS Martin Enserink(Europe), Mara Hvistendahl (Asia) SR. CORRESPONDENTS Jeffrey Mervis, Elizabeth Pennisi NEWS WRITERS Adrian Cho, Jon Cohen, Jennifer Couzin-Frankel, Carolyn Gramling, Eric Hand, Jocelyn Kaiser, Kelly Servick, Robert F. Service, Erik Stokstad, Emily Underwood INTERNS David Shultz, Jia You CONTRIBUTING CORRESPONDENTS Pallava Bagla(South Asia), Michael Balter(Paris), John Bohannon, Ann Gibbons, Sam Kean, Richard A. Kerr, Eli Kintisch, Kai Kupferschmidt(Berlin), Andrew Lawler, Christina Larson(Beijing), Mitch Leslie, Charles C. Mann, Eliot Marshall, Virginia Morell, Dennis Normile(Tokyo), Heather Pringle, Tania Rabesandratana(Brussels), Gretchen Vogel(Berlin), Lizzie Wade(Mexico City) CAREERS Jim Austin(Editor), Donisha Adams COPY EDITORS Kara Estelle, Nora Kelly, Jennifer Levin ADMINISTRATIVE SUPPORT Scherraine Mack

## Executive Publisher Alan I. Leshner

Publisher Kent R. Anderson Chief Digital Media Officer Rob Covey

BUSINESS OPERATIONS AND ADMINISTRATION DIRECTOR Deborah Rivera-Wienhold BUSINESS SYSTEMS AND FINANCIAL ANALYSIS DIRECTOR Randy Yi MANAGER OF FULFILLMENT SYSTEMS Neal Hawkins SYSTEMS ANALYST Nicole Mehmedovich ASSISTANT DIRECTOR, BUSINESS OPERATIONS Eric Knott MANAGER, BUSINESS OPERATIONS Jessica Tierney BUSINESS ANALYSTS Cory Lipman, Cooper Tilton, Celeste Troxler FINANCIAL ANALYST Jeremy Clay RIGHTS AND PERMISSIONS ASSISTANT DIRECTOR Emilie David PERMISSIONS ASSOCIATE Elizabeth Sandler RIGHTS, CONTRACTS, AND LICENSING ASSOCIATE Lili Kiser

MARKETING DIRECTOR Ian King MARKETING MANAGER Julianne Wielga MARKETING ASSOCIATE Elizabeth Sattler SR. MARKETING EXECUTIVE Jennifer Reeves SR. ART ASSOCIATE, PROJECT MANAGER Tzeitel Sorrosa ART ASSOCIATE Seil Lee ASSISTANT COMMERCIAL EDITOR Selby Frame MARKETING PROJECT MANAGER Angelissa McArthur SR. WRITER Bill Zimmer PROGRAM DIRECTOR, AAAS MEMBER CENTRAL Peggy Mihelich FULFILLMENT SYSTEMS AND OPERATIONS membership@aaas.org MANAGER, MEMBER SERVICES Pat Butler SPECIALISTS LaToya Casteel, Javia Flemmings, Latasha Russell MANAGER, DATA ENTRY Mickie Napoleoni DATA ENTRY SPECIALISTS JJ Regan, Jaimee Wise, Fiona Giblin

DIRECTOR, SITE LICENSING Tom Ryan DIRECTOR, CORPORATE RELATIONS Eileen Bernadette Moran SR. PUBLISHER RELATIONS SPECIALIST Kiki Forsythe PUBLISHER RELATIONS MANAGER Catherine Holland PUBLISHER RELATIONS, EASTERN REGION Keith Layson PUBLISHER RELATIONS, WESTERN REGION Ryan Rexroth MANAGER, SITE LICENSE OPERATIONS Iquo Edim FULFILLMENT ANALYST Lana Guz ASSOCIATE DIRECTOR, MARKETING Christina Schlecht MARKETING ASSOCIATES Thomas Landreth, Minah Kim

DIRECTOR OF WEB TECHNOLOGIES Ahmed Khadr SR. DEVELOPER Chris Coleman DEVELOPERS Dan Berger, Jimmy Marks SR. PROJECT MANAGER Trista Smith SYSTEMS ENGINEER Luke Johnson PRODUCT MANAGER Walter Jones

CREATIVE DIRECTOR, MULTIMEDIA Martyn Green DIRECTOR OF ANALYTICS Enrique Gonzales SR. WEB PRODUCER Sarah Crespi WEB PRODUCER Alison Crawford VIDEO PRODUCER Nguyen Nguyen SOCIAL MEDIA PRODUCER Meghna Sachdev

DIRECTOR OF OPERATIONS PRINT AND ONLINE Elizabeth Harman PRINT PRODUCTION DIRECTOR Wendy K. Shank PREFLIGHT MANAGER Marcus Spiegler ASSISTANT MANAGER ONLINE Lisa Stanford ASSISTANT MANAGER PRINT Rebecca Doshi SR. SPECIALISTS Steve Forrester, Jason Hillman, Antoinette Hodal, Tara Kelly, Anthony Rosen SPECIALISTS Jacob Hedrick, Nichole Johnston, Lori Murphy, Kimberley Oster

DESIGN DIRECTOR Beth Rakouskas ASSOCIATE ART DIRECTOR Laura Creveling SR. ILLUSTRATORS Chris Bickel, Katharine Sutliff ILLUSTRATOR Valerie Altounian SR. ART ASSOCIATES Holly Bishop, Preston Huey ART ASSOCIATES Kay Engman, Garvin Grullón, Chrystal Smith SR. PHOTO EDITOR William Douthitt PHOTO EDITOR Leslie Blizard

DIRECTOR, GLOBAL COLLABORATION, CUSTOM PUBLICATIONS, ADVERTISING Bill Moran EDITOR, CUSTOM PUBLISHING Sean Sanders: 202-326-6430 ASSISTANT EDITOR, CUSTOM PUBLISHING Tianna Hicklin: 202-326-6463 ADVERTISING MARKETING MANAGER Justin Sawyers: 202-326-7061 science\_advertising@aaas.org ADVERTISING MARKETING ASSOCIATE Javia Flemmings ADVERTISING SUPPORT MANAGER Karen Foote: 202-326-6740 ADVERTISING PRODUCTION OPERATIONS MANAGER Deborah Tompkins SR. PRODUCTION SPECIALIST/GRAPHIC DESIGNER Amy Hardcastle PRODUCTION SPECIALIST Yuse Lajiminmuhup SR. TRAFFIC ASSOCIATE Christine Hall SALES COORDINATOR Shirley Young ASSOCIATE DIRECTOR, COLLABORATION, CUSTOM PUBLICATIONS/CHINA/TAIWAN/KOREA/SINGAPORE Ruolei Wu: +86-186 0822 9345, ruw@aaas.org COLLABORATION/CUSTOM PUBLICATIONS/JAPAN Adarsh Sandhu + 81532-81-5142 asandhu@aaas.org EAST COAST/E. CANADA Laurie Faraday: 508-747-9395, FAX 617-507-8189 WEST COAST/W. CANADA Lynne Stickrod: 415-931-9782, FAX 415-520-6940 MIDWEST Jeffrey Dembski: 847-498-4520 x3005, Steven LeRoy: 847-498-4520 x3006 UK EUROPE/ASIA Roger Goncalves: TEL/FAX +41 43 243 1358 JAPAN Katsuyoshi Fukamizu(Tokyo): +81-3-3219-5777 kfukamizu@aaas.org CHINA/TAIWAN Ruolei Wu: +86-10682-9345

WORLDWIDE ASSOCIATE DIRECTOR OF SCIENCE CAREERS Tracy Holmes: +44 (0) 1223 326525, FAX +44 (0) 1223 326532 tholmes@science-int.co.uk CLASSIFIED advertise@sciencecareers.org U.S. SALES Tina Burks: 202-326-6577 Nancy Toema: 202-326-6578 SALES ADMINISTRATOR Marci Gallun EUROPE/ROW SALES Axel Gesatzki, Sarah LeMay SALES ASSISTANT Kelly Grace JAPAN Hiroyuki Mashiki(Kyoto): +81-75-823-1109 hmashiki@aaas.org CHINA/TAIWAN Ruolei Wu: +86-186 0082 9345 ruw@aaas.org MARKETING MANAGER Allison Pritchard MARKETING ASSOCIATE Aimee Aponte

AAAS BOARD OF DIRECTORS RETIRING PRESIDENT, CHAIR Phillip A. Sharp PRESIDENT Gerald R. Fink PRESIDENT-ELECT Geraldine (Geri) Richmond TREASURER David Evans Shaw CHIEF EXECUTIVE OFFICER Alan I. Leshner BOARD PRESIDENT L. Baggler, May R. Berenbaum, Carlos J. Bustamante, Claire M. Fraser, Laura H. Greene, Elizabeth Loftus, Raymond Orbach, Inder M. Verma

SUBSCRIPTION SERVICES For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005 INSTITUTIONAL SITE LICENSES 202-326-6755 REPRINTS: Author Inquiries 800-635-7181 COMMERCIAL INQUIRIES 803-359-4578 PERMISSIONS 202-326-6765, permissions@aaas.org AAAS Member Services 202-326-6417 or http://membercentral.aaas.org/discourses

Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority or conflicting points of view. Accordingly, all articles published in Science—including editorials, news and comment, and books reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

INFORMATION FOR AUTHORS See pages 680 and 681 of the 7 February 2014 issue or access [www.sciencemag.org/about/authors](http://www.sciencemag.org/about/authors)

## SENIOR EDITORIAL BOARD

A. Paul Alivisatos, Lawrence Berkeley Nat'l Laboratory, Ernst Fehr, U. of Zürich  
Susan M. Rosenberg, Baylor College of Medicine, Michael S. Turner, U. of Chicago

## BOARD OF REVIEWING EDITORS (Statistics board members indicated with \$)

Adriano Aguzzi, U. Hospital Zürich  
Takuzo Aida, U. of Tokyo  
Leslie Aiello, Wenner-Gren Foundation  
Judith Allen, U. of Edinburgh  
Sonia Altizer, U. of Georgia  
Virginia Armbrust, U. of Washington  
Sebastian Amigorena, Institut Curie  
Kathryn Anderson, Memorial Sloan-Kettering Cancer Center  
Peter Andolfatto, Princeton U.  
Meinrat O. Andreae, Max-Planck Inst. Mainz  
Paola Ariotta, Harvard U.  
Johan Auwerx, EPFL  
David Awschalom, U. of Chicago  
Jordi Bascompte, Estación Biológica de Doñana CSIC  
Facundo Batista, London Research Inst.  
Ray H. Baughman, U. of Texas, Dallas  
David Baum, U. of Wisconsin  
Kamran Behnia, ESPCI-ParisTech  
Yasmine Belkaid, NIAID, NIH  
Philip Benfey, Duke U.  
Stephen J. Benkovic, Penn State U.  
Carlo Beenakker, Leiden U.  
Gabriele Bergers, U. of California, San Francisco  
Christophe Bernard, Aix-Marseille U.  
Bradley Bernstein, Massachusetts General Hospital  
Peer Bork, EMBL  
Bernard Bourdon, Ecole Normale Supérieure de Lyon  
Chris Bowler, Ecole Normale Supérieure  
Ian Boyd, U. of St. Andrews  
Emily Brodsky, U. of California, Santa Cruz  
Ron Brookmeyer, U. of California Los Angeles (\$)   
Christian Büchel, U. Hamburg-Eppendorf  
Joseph A. Burns, Cornell U.  
Gyorgy Buzsáki, New York U. School of Medicine  
Blanche Capel, Duke U.  
Mats Carlsson, U. of Oslo  
David Clapham, Children's Hospital Boston  
David Clary, U. of Oxford  
Joel Cohen, Rockefeller U., Columbia U.  
Jonathan D. Cohen, Princeton U.  
James Collins, Boston U.  
Robert Cook-Deegan, Duke U.  
Alan Cowman, Walter & Eliza Hall Inst.  
Robert H. Crabtree, Yale U.  
Roberta Croce, Vrije Universiteit  
Janet Currie, Princeton U.  
Jeff L. Dangl, U. of North Carolina  
Tom Daniel, U. of Washington  
Frans de Waal, Emory U.  
Stanislav Dehaene, Collège de France  
Robert Desimone, MIT  
Claude Despland, New York U.  
Ap Dijksterhuis, Radboud U. of Nijmegen  
Dennis Discher, U. of Pennsylvania  
Gerald W. Dorn II, Washington U. School of Medicine  
Jennifer A. Doudna, U. of California, Berkeley  
Bruce Dunn, U. of California, Los Angeles  
Christopher Dye, WHO  
Todd Ehlers, U. of Tuebingen  
David Ehrhardt, Carnegie Inst. of Washington  
Tim Elston, U. of North Carolina at Chapel Hill  
Gerhard Ertl, Fritz-Haber-Institut, Berlin  
Barry Everitt, U. of Cambridge  
Ernst Fehr, U. of Zurich  
Anne C. Ferguson-Smith, U. of Cambridge  
Michael Feuer, The George Washington U.  
Kate Fitzgerald, U. of Massachusetts  
Peter Fratzl, Max-Planck Inst.  
Elaine Fuchs, Rockefeller U.  
Daniel Geschwind, UCLA  
Andrew Gewirth, U. of Illinois  
Karl-Heinz Glassmeier, TU Braunschweig  
Ramon Gonzalez, Rice U.  
Julia R. Greer, Caltech  
Elizabeth Grove, U. of Chicago  
Kip Guy, St. Jude's Children's Research Hospital  
Taekjip Ha, U. of Illinois at Urbana-Champaign  
Christian Haass, Ludwig Maximilians U.  
Steven Hahn, Fred Hutchinson Cancer Research Center  
Michael Hasselmo, Boston U.  
Martin Heimann, Max-Planck Inst. Jena  
Yka Helariutta, U. of Cambridge  
James A. Hendler, Rensselaer Polytechnic Inst.  
Janet G. Hering, Swiss Fed. Inst. of Aquatic Science & Technology  
Michael E. Himmel, National Renewable Energy Lab.  
Kai-Uwe Hinrichs, U. of Bremen  
Kei Hirose, Tokyo Inst. of Technology  
David Hodell, U. of Cambridge  
David Holden, Imperial College  
Lora Hooper, UT Southwestern Medical Ctr. at Dallas  
Raymond Huey, U. of Washington  
Steven Jacobsen, U. of California, Los Angeles  
Kai Johnsson, EPFL Lausanne  
Peter Jonas, Inst. of Science & Technology (IST) Austria  
Matt Kaeblerlein, U. of Washington  
William Kaelin Jr., Dana-Farber Cancer Inst.  
Daniel Kahne, Harvard U.  
Daniel Kammen, U. of California, Berkeley  
Masashi Kawasaki, U. of Tokyo  
Joel Kingsolver, U. of North Carolina at Chapel Hill  
Robert Kingston, Harvard Medical School  
Eitonne Koechlin, Ecole Normale Supérieure  
Alexander Kolodkin, Johns Hopkins U.  
Roberto Kolter, Harvard Medical School  
Alberto R. Kornblith, U. of Buenos Aires  
Leonid Kruglyak, UCLA  
Thomas Langer, U. of Cologne  
Mitchell A. Lazar, U. of Pennsylvania  
David Lazer, Harvard U.  
Thomas Lecuit, IBM  
Virginia Lee, U. of Pennsylvania  
Stanley Lemon, U. of North Carolina at Chapel Hill  
Ottoline Leyser, Cambridge U.  
Marcia C. Linn, U. of California, Berkeley  
Jianguo Liu, Michigan State U.  
Luis Liz-Marzan, CIC bioGUNE  
Jonathan Losos, Harvard U.  
Ke Lu, Chinese Acad. of Sciences  
Christian Lüscher, U. of Geneva  
Laura Machesky, CRUK Beatson Inst. for Cancer Research  
Anne Magurran, U. of St. Andrews  
Oscar Marin, CSIC & U. Miguel Hernández  
Charles Marshall, U. of California, Berkeley  
C. Robertson McClung, Dartmouth College  
Graham Medley, U. of Warwick  
Yasushi Miyashita, U. of Tokyo  
Richard Morris, U. of Edinburgh  
Alison Moutsier-Reif, NC State U. (\$)   
Sean Munro, MRC Lab. of Molecular Biology  
Thomas Murray, The Hastings Center  
James Nelson, Stanford U. School of Med.  
Karen Nelson, U. of California, Berkeley  
Daniel Neumark, U. of California, Berkeley  
Timothy W. Nilsen, Case Western Reserve U.  
Pär Nordlund, Karolinska Inst.  
Helga Nowotny, European Research Advisory Board  
Ben Olken, MIT  
Joe Orenstein, U. of California  
Berkeley & Lawrence Berkeley National Lab  
Harry Orr, U. of Minnesota  
Andrew Oswald, U. of Warwick  
Steve Palumbi, Stanford U.  
Jane Parker, Max-Planck Inst. of Plant Breeding Research  
Giovanni Parmigiani, Dana-Farber Cancer Inst. (\$)   
Donald R. Paul, U. of Texas, Austin  
John H. J. Petrini, Memorial Sloan-Kettering Cancer Center  
Joshua Plotkin, U. of Pennsylvania  
Albert Polman, FOM Institute AMOLF  
Philippe Poulin, CNRS  
David Randall, Colorado State U.  
Colin Renfrew, U. of Cambridge  
Felix Rey, Institut Pasteur  
Trevor Robbins, U. of Cambridge  
Jim Roberts, Fred Hutchinson Cancer Research Ctr.  
Barbara A. Romanowicz, U. of California, Berkeley  
Jens Rostrup-Nielsen, Haldor Topsøe  
Mike Ryan, U. of Texas, Austin  
Mitsunori Saitou, Kyoto U.  
Shimon Sakaguchi, Kyoto U.  
Miquel Salmeron, Lawrence Berkeley National Lab  
Jürgen Sandkühner, Medical U. of Vienna  
Alexander Schier, Harvard U.  
Randy Seeley, U. of Cincinnati  
Vladimir Shaleev, Purdue U.  
Robert Siliciano, Johns Hopkins School of Medicine  
Joseph Silk, Institut d'Astrophysique de Paris  
Denis Simon, Arizona State U.  
Alison Smith, John Innes Centre  
Richard Smith, U. of North Carolina (\$)   
John Speakman, U. of Aberdeen  
Allan C. Spradling, Carnegie Institution of Washington  
Jonathan Sprent, Garvan Inst. of Medical Research  
Eric Steig, U. of Washington  
Paula Stephan, Georgia State U. and National Bureau of Economic Research  
Molly Stevens, Imperial College London  
V. S. Subrahmanian, U. of Maryland  
Ira Tabas, Columbia U.  
Sarah Teichmann, Cambridge U.  
John Thomas, North Carolina State U.  
Shubha Tole, Tata Institute of Fundamental Research  
Christopher Tyler-Smith, The Wellcome Trust  
Sanger Inst.  
Herbert Virsik, Washington U.  
Bern Vogelstein, Johns Hopkins U.  
Cynthia Volkert, U. of Göttingen  
Douglas Wallace, Dalhousie U.  
David Wallace, Weizmann Inst. of Science  
Ian Walsley, U. of Oxford  
David A. Wardle, Swedish U. of Agric. Sciences  
David Waxman, Fudan U.  
Jonathan Weissman, U. of California, San Francisco  
Chris Wickle, U. of Missouri (\$)   
Ian A. Wilson, The Scripps Res. Inst. (\$)   
Timothy D. Wilson, U. of Virginia  
Rosemary Wyse, Johns Hopkins U.  
Jan Zaenen, Leiden U.  
Kenneth Zaret, U. of Pennsylvania School of Medicine  
Jonathan Zehr, U. of California, Santa Cruz  
Len Zon, Children's Hospital Boston  
Maria Zuber, MIT

## BOOK REVIEW BOARD

David Bloom, Harvard U., Samuel Bowring, MIT, Angela Creager, Princeton U.,  
Richard Swedner, U. of Chicago, Ed Wasserman, DuPont

# Out of sight, out of mind

The 2014 mid-term elections are over, and at least 62 new members of the U.S. Congress will be taking office on 3 January 2015. Many will be accompanied by new staff, not all of them familiar with science and science policy. Compounding the ambiguity that always comes with new legislators is that the leadership of the Senate has changed from Democratic to Republican. It is as yet unknown where the new leaders and committee chairs will stand on science policy issues, but so far, science has not been declared among the priorities put forward by either party. All of this is taking place just as many of science's long-time legislative champions are retiring. Overall research and development spending has already fallen 16% in inflation-adjusted dollars from fiscal year (FY) 2010 to the FY 2015 budget request. The federal government's investment in science and technology now stands at roughly 0.78% of the economy, the lowest point in 50 years. This is the right time for scientists to introduce themselves to their new legislators, congratulate them on their election, offer their services if needed on policy matters that would benefit from science input, and emphasize the importance of science to the country.

It is imperative that scientists develop relationships with new members of Congress and their staff. There will be another Ebola, earthquake, flood, or oil spill, and a host of other science-related challenges. Newly elected members will soon be asked to vote on these kinds of issues. They will need information quickly to make knowledge-based decisions. If scientists do not act now to cultivate new, well-informed champions for science, other interests will quickly eclipse science in vying for attention and loyalty. This cultivation is particularly important now, at a time of intense scrutiny of the federal government's investment in all programs.

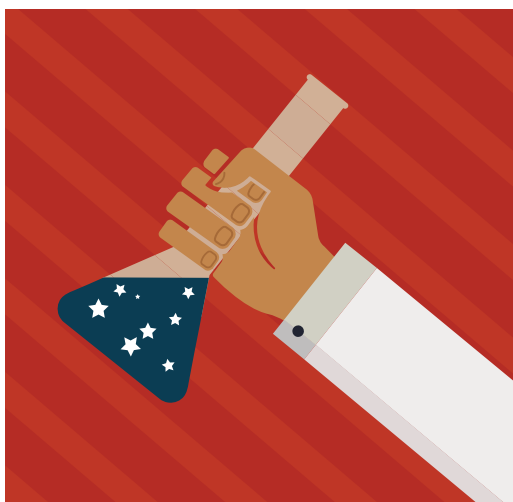
Thinking more strategically, the scientific community should take a page from the playbook of those who successfully run for public office. Unlike most scientists, elected officials are practiced at demonstrating accountability to the public. If elected officials do not regularly say and convey, "I work for you, and this is what I've done for you lately," they will not be back for a next term. If a member of Congress is effectively out of sight, unconnected to their district, he or she will not be reelected.

Politicians know how to be visible and accountable to the taxpayer; scientists, not so much. Surveys commissioned by Research!America show that scientists are essentially invisible in the United States, with only 30% of Americans able to name a living scientist and 44% able to name a place where research is conducted.\* It's all too easy for members of the science community to hunker down and stay out of policy-makers' view, especially given that the culture of science often discourages, rather than embraces, engagement with the nonscience public. It is also easy to be intimidated by outreach to policy-makers, not least but also not only because little or no training in dealing with policy-makers and with

public communication is provided in graduate science education, and little encouragement or recognition is on offer thereafter.

It is past time to change counterproductive habits, and there are resources with tactical advice on how to do so (such as the AAAS guide to working with Congress†). If the scientific community stays out of sight much longer, the issues of science will be even further out of the public's and policy-makers' minds and then, if the priority of science sinks further, we will all be out of luck!

— Mary Woolley and Alan I. Leshner



***"...if the priority of science sinks further, we will all be out of luck!"***



*Mary Woolley is the president and chief executive officer of Research!America. E-mail: mwoolley@researchamerica.org*



*Alan I. Leshner is the chief executive officer of the American Association for the Advancement of Science and executive publisher of Science*

\*[www.researchamerica.org](http://www.researchamerica.org)

†[www.aaas.org/news/top-ten-rules-working-congress](http://www.aaas.org/news/top-ten-rules-working-congress)



“The Committee’s ... substituting [its] judgment for the expertise of scientists.”

The Association of American Universities, on a review of National Science Foundation grants by the science committee of the U.S. House of Representatives.

## IN BRIEF

### A pregnant mare and fetus, fossilized



The fossilized remains of a pregnant mare and fetus (white circle) offer a glimpse into the reproduction of *E. messelensis*.

**F**orty-seven million years ago, a pregnant mare and its unborn foal lost their lives—perhaps chased into a lake, where they drowned. The fossilized remains of the pair, from the early horse species *Eurohippus messelensis*, were discovered in the Messel Pit, a former coal and oil shale mine near Frankfurt, Germany, that is famous for its exquisitely preserved fossils. Like other early horses, the mare was small, about the size of a fox terrier, says Jens Franzen, a paleontologist at the Senckenberg Research Institute and Natural History Museum in Frankfurt, who presented the fossil last week at the Society of Vertebrate Paleontology annual meeting in Berlin. Not only were most of the bones of the mare and her fetus intact, but scientists can also detect the placenta, which was not fossilized directly: It is visible as a dark shadow left by bacteria that consumed the tissue and were then fossilized. The position of the foal suggests that it was close to mature, and suggests that ancient horses gave birth in a similar way to their modern cousins.

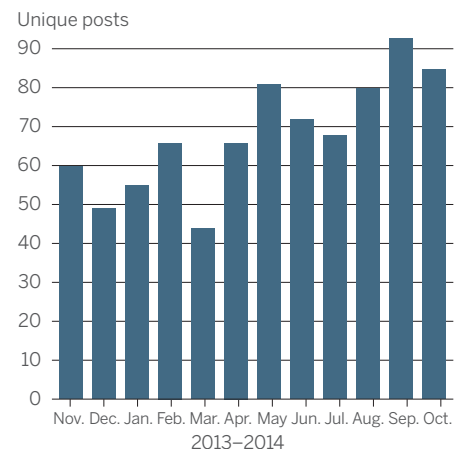
## AROUND THE WORLD

### USDA OKs engineered potatoes

**WASHINGTON, D.C.** | The U.S. Department of Agriculture has approved commercial planting of a potato that is genetically engineered, but not transgenic. Created by the J. R. Simplot Company in Boise, the potatoes resist bruising and, when fried, contain up to 80% less of the suspected carcinogen acrylamide than comparable varieties. Simplot created its new potatoes by adding and rearranging DNA from wild and domesticated varieties; thanks to RNA interference, two genes do not make the enzymes that contribute to bruising and the creation of acrylamide. In the late 1990s, Monsanto created a pest-resistant potato by adding in a bacterial gene, but took it off the market due to consumer opposition. Simplot expects that several thousand hectares will be planted with the new varieties next year, mostly to supply supermarkets.

### BioRxiv at 1 year

**COLD SPRING HARBOR, NEW YORK** | A year after its launch, a new preprint server for life sciences is off to a healthy start, its creators say. BioRxiv, a free site sponsored by the Cold Spring Harbor Laboratory (CSHL), has attracted more than 800 papers (see graph). Submissions span many disciplines—such as cell biology



BioRxiv has attracted a growing stream of papers, now totaling 824 new manuscripts.

Adult mosquito  
emerges from its  
watery nursery.



## Craving human blood is in mosquito's DNA

**R**esearchers have isolated a gene in the mosquito *Aedes aegypti* that helps explain why the species has such a yen for human blood. *A. aegypti*, which transmits dengue and other viruses to humans, can interbreed with a close, nonhuman-biting, forest-dwelling relative in Africa. Carolyn McBride, an evolutionary neurobiologist at Princeton University, and her colleagues compared the two subspecies, finding that human-biting “domestics” have distinct versions of a gene called *Or4* and make a lot more of this odor-sensing protein than the forest mosquitoes. McBride and her colleagues report this week in *Nature*. The domestic mosquito homes in on an odor component called sulcatone, which humans produce in abundance. Although there are likely to be other factors involved in human preference, “it seems the one identified is a, if not *the*, major factor,” comments Jeffrey Powell, an evolutionary geneticist at Yale University. <http://scim.ag/mosqDNA>

and cancer research—in which preprint sharing hadn't been routine. About 28% of authors, who come from 44 countries, have revised their papers, presumably after getting feedback from readers, says John Inglis, executive director of CSHL Press. For scientists who might worry that posting a preprint will jeopardize its chances at a journal, Inglis points out that one-fourth of bioRxiv's papers have later appeared in journals including *Science*, *Nature*, and *Cell*. <http://scim.ag/bioRxiv1yr>

## Historic Lick Observatory saved

**SAN JOSE, CALIFORNIA** | The Lick Observatory has gained a new lease on life. The first permanent mountaintop observatory in the world when it opened in 1888, Lick was targeted for closing in 2013 by University of California (UC) administrators faced with shrinking funding. A “Save Lick” campaign spearheaded by UC astronomers and amateur stargazers

ensued. Then, in a 29 October letter, the administrators said they had scrapped the plan, noting that University of California Observatories (UCO), which manages the observatory program for the university system, had convinced them that UCO could pay for a slimmed-down program without sacrificing other priorities. But the observatory's financial future remains tricky, with a spartan \$1.5-million-a-year budget and shrinking staff. <http://scim.ag/LickObserv>

## Ebola beds still needed, says WHO

**GENEVA, SWITZERLAND** | Some once-overwhelmed Ebola treatment centers now have empty beds because the epidemic has waned in certain areas, particularly Monrovia—but the World Health Organization (WHO) warns that a serious bed shortage still exists elsewhere in the country, as well as in hard-hit Sierra Leone and Guinea. According to WHO's 5 November Ebola “situation report,” only

## BY THE NUMBERS

999

Number of unique proteins produced by human sperm—compared with the human brain's 318—according to the latest version of the online Human Protein Atlas, launched last week.

25%

Estimated increase in pollen production—and perhaps allergy suffering—as atmospheric CO<sub>2</sub> increases from the current 400 ppm to 450 ppm, suggests a study in *PLOS ONE*.

\$1.2  
billion

Amount that Google will pay NASA over 60 years to lease its historic Moffett airfield in northern California. Google plans to spend \$200 million to convert the 83-year-old hangar into a research center for space, aviation, and robotics technologies.

22% of a planned 4707 beds are “in operation.” A key problem, the report notes, is a shortage of medical teams, not the beds themselves. All told, the three countries had only 10 of 38 needed medical teams.

## NEWSMAKERS

### Polio pioneer heads India science

Science now has a more potent voice in India's government. Prime Minister Narendra Modi appointed physician **Harsh Vardhan** on 9 November as minister for science and technology and earth sciences and elevated him to full Cabinet rank. Vardhan, 59, had served as India's health minister since Modi formed his government in May; he is known for his pioneering role in the eradication of polio from India, which earlier this year was declared free of wild poliovirus. As health minister he also promoted Ayurveda, the traditional Indian system of medicine.





After his acquittal, Giulio Selvaggi (left) talks to Bernardo De Bernardinis, who remains convicted.

the notion that the earlier, smaller tremors were a good thing because they discharged energy. Witnesses at the original trial said that idea—which most scientists regard as incorrect—led their relatives to remain indoors.

The defense lawyers claimed that this idea could not have been endorsed by the commission as a whole, as Billi argued, because it was stated publicly only by De Bernardinis, in an interview before the commission's meeting. "No one said, 'Stay at home because there is a discharge of energy,'" said Marcello Melandri, the lawyer for Enzo Boschi, former head of Italy's National Institute of Geophysics and Volcanology. But prosecutor Romolo Como pointed out that none of the participants had replied when asked about the energy-discharge idea during the meeting. "Why didn't anyone object to this?" he demanded.

In the new verdict, which must be explained in writing within 3 months, a panel of three judges headed by Fabrizia Francabandera concluded that only in De Bernardinis's case could a link be proven between the expert's words and the actions of some of the victims. De Bernardinis said that, notwithstanding his renewed conviction, he could face "God and men" with a clear conscience. Some L'Aquila citizens afterward expressed indignation at what they saw as a further betrayal on the part of the Italian state. Lawyers for relatives of several of the victims say they will appeal the verdict in the Supreme Court of Cassation in Rome.

Boschi told *Science* that he felt unwell during the final hearing and that he only realized he had been acquitted a quarter of an hour after the verdict was read out. He says that "there is nothing to celebrate," but that he is "relieved" and hopes eventually to go and speak with people in L'Aquila "without lawyers."

Willy Aspinall, a risk expert at the University of Bristol in the United Kingdom, is relieved as well. But he believes the trial should serve as a warning. The L'Aquila risk analysis "wasn't done properly, competently, and neutrally," says Aspinall, adding that the brief duration of the commission's meeting—it lasted about three quarters of an hour—showed the seven were "taking a bit of a gamble that they could serve their political bosses and that would be the end of it." He argues that scientists should be "much more cautious about how they give advice, to put it in writing rather than just verbally." ■

*Edwin Cartlidge is a science writer in Rome.*

## IN DEPTH

### SEISMOLOGY

# Relief greets acquittals in Italy earthquake trial

But some still fault scientists for hasty risk assessment

By Edwin Cartlidge, in L'Aquila, Italy

For families of the victims, it came as a stunning, painful reversal—"an earthquake after the earthquake," a lawyer for several of them said. But many scientists are relieved by the latest development in a 5-year legal saga that triggered heated debates in Italy and unnerved earth scientists around the world.

On Monday, an appeals court acquitted six experts convicted of manslaughter and sentenced to 6 years in prison in 2012 for advice they gave ahead of a 2009 earthquake that killed 309 people in this provincial town. Only one of the seven originally found guilty remains convicted: Bernardo De Bernardinis, who in 2009 was deputy head of Italy's Civil Protection Department and who has been given a 2-year suspended jail sentence, pending any further appeals. "The L'Aquila trial was a false prosecution of individuals for what was a failure of a national risk communication system," says Thomas Jordan, an earth scientist at the University of Southern California in Los Angeles. "Judicial sanity prevailed."

Only 6 days before the fatal quake struck, the experts had attended a meeting of the National Commission for the Forecast and Prevention of Major Risks. The meeting, held

on 31 March 2009, had been organized to evaluate the threat posed by a series of small and medium-sized tremors that had been shaking L'Aquila for several months. In the 2012 trial, Judge Marco Billi ruled that the commission members carried out a "superficial, approximate and generic" risk analysis, and that they made a number of reassuring statements that led 29 of the quake's victims to remain indoors at the time of the disaster, even though two moderate tremors had struck several hours beforehand.

The tough sentence Billi handed down shocked the Italian research community. It led many to argue that science itself had been found guilty, even though the judge emphasized that he had not convicted the experts for having failed to predict the earthquake—which science can't do—but rather for having failed to carry out their legally binding duties as "public officials." He said the experts hadn't analyzed a series of factors indicating a heightened seismic risk, including the fact that previous quakes that destroyed L'Aquila were preceded by smaller tremors.

Lawyers for the convicted experts took aim at both the alleged negligence of their clients and the claim that their statements swayed people's decision to stay indoors, in buildings that later collapsed. They disagreed that the panel could be blamed for

## INFECTIOUS DISEASES

# Just one poliovirus left to go?

Vaccinators may have wiped out the second of three strains of the virus

By Leslie Roberts

**T**he long-running global polio eradication effort, years behind schedule and buffeted by setbacks, got a shot of good news this week. Its leaders declared the effort has likely eradicated the second of the three varieties of the virus.

The feat, if confirmed, would mark just the third time that a distinct human pathogen has been wiped out through immunization, after smallpox and one other polio strain, the scientists write. “This is a very big deal ... that puts us a lot closer to eradication” of all three polioviruses, says Chris Maher, the World Health Organization’s manager for polio eradication and emergency response in the eastern Mediterranean region, who is not on the new paper. “Two of the three viruses [are] down, and only one to go,” he adds. The news “bodes well” for the Global Polio Eradication Initiative (GPEI), agrees Walter Orenstein, an infectious disease expert at Emory University School of Medicine in Atlanta.

The announcement on 13 November in the *Morbidity and Mortality Weekly Report* (MMWR) comes with caveats. The paper is titled “Possible Eradication of Wild Poliovirus Type 3—Worldwide, 2012.” “Given all the bumps in the road, we are trying to be ... cautious,” says Stephen Cochi, a co-author and one of the lead scientists on polio eradication at the U.S. Centers for Disease Control and Prevention (CDC) in Atlanta. But “we believe the evidence is quite good,” says first author Olen Kew, a CDC virologist who has tracked the global comings and goings of the virus since 1985.

Poliovirus comes in three varieties, or serotypes, each genetically distinct and with its own personality. Wild-type 2, for instance, is the most transmissible, whereas wild-type 3 tends to stay put. Wild-type 3 is the “most insidious of the viruses,” Maher says, because it causes paralysis in only about one in every 1000 people it infects, compared with one in 200 for wild-type 1, and so can circulate undetected. The trivalent oral polio vaccine (OPV) that is a workhorse of the campaign is essentially three vaccines in one, and it works better against some serotypes than others.

Wild-type 2 was the first to disappear globally, in 1999, although it persists in what is known as a vaccine-derived form,

created when the attenuated virus used in OPV reverts to its virulent, transmissible form. For reasons that are not entirely clear, in most of the countries that have already eradicated polio, wild-type 3 “checked out” next, a number of years later, followed closely by wild-type 1, Kew says. In Vietnam, for instance, type 3 vanished in 1993, and type 1 in 1994. India, the most recent country to eradicate polio, last saw type 3 virus in 2010; type 1 held on just one more year. The eradicationists fervently hope the pattern will hold globally, and in the past few years, GPEI has gone on a war footing, with more targeted vaccines and a surge of public health workers into infected countries.

The last two sightings of wild-type 3 occurred in Nigeria in 2012, where an 11-month-old was paralyzed in the northern state of Yobe on 10 November. The last positive environmental sample was collected from sewage

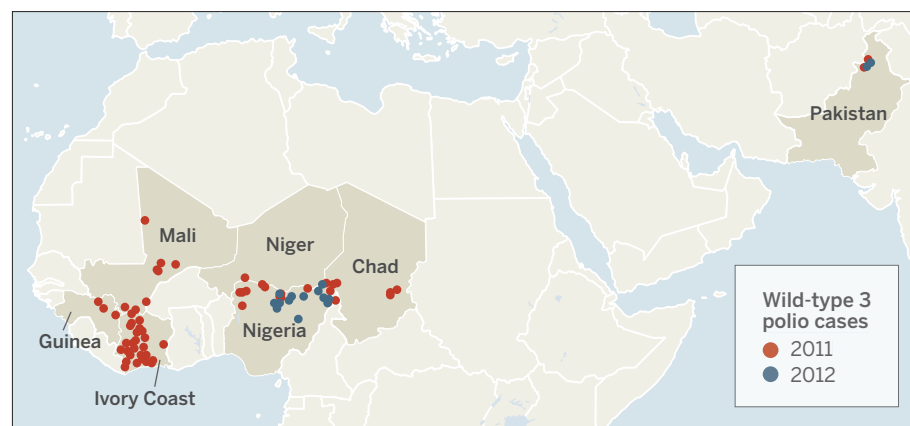
virus is to blame, GPEI now tests environmental samples of sewage water for traces of the virus. This “would pick up the virus even without cases,” Kew says, as happened recently in Israel with wild-type 1 (*Science*, 8 November 2013, p. 679).

Genetic evidence also points to extinction, Kew says. Wild-type 3 did not suddenly vanish. “It is not a binary event, now you have it, now you don’t,” Kew says. But over the years, the genetic diversity of this serotype declined, from 17 distinct genotypes in 1988 to just two in 2012.

The disappearance of type 3 from Nigeria adds to the encouraging news from that country. It has not seen a trace of wild poliovirus in 3 months and has had only six cases this year. “I am encouraged we will see a polio-free Africa by the end of the year,” Cochi says. “The one lingering uncertainty is Borno, the northeastern state where [the terrorist group] Boko Haram is most active.”

## A virus fades out

Polio cases caused by wild-type 3 dropped from 67 in 2011 to 21 in 2012, then vanished.



in Lagos the following day. Despite scouring the genetic evidence for wild poliovirus type 3 ever since, “we can’t find it anywhere,” Kew says. Although encouraged, Orenstein would like one more year of intense surveillance to be sure. “In my opinion, it is premature to celebrate type 3’s eradication.”

The CDC scientists say they made the announcement now because 2 years of intense surveillance have passed since the last sighting. In addition to investigating stool samples from every case of acute flaccid paralysis worldwide to see whether polio-

The MMWR announcement may also put even more pressure on Pakistan, which has a runaway epidemic. The country now accounts for 85% of all polio cases in the world and constantly reinfects neighboring Afghanistan. With its ballooning case count—more than 235 this year—Pakistan is fast becoming a global “pariah,” Cochi says. If GPEI’s intensified strategy works in Nigeria, with its troubles and insecurity, there is no reason it can’t work in Pakistan, he says—if the government finally gets serious about polio eradication. ■



## U.S. ELECTIONS

# For science, impact of Republican win may be muted

Congress could still act on some research-related issues despite worsening relations with the White House

By Jeffrey Mervis and David Malakoff

In January, control of the U.S. Senate's environment committee will shift from a legislator who believes "climate change is a catastrophe that is unfolding before our eyes" to one who thinks "anthropogenic climate change is the world's greatest hoax."

That stark ideological reversal, with Senator Barbara Boxer (D-CA) ceding the reins to Senator James Inhofe (R-OK), is one result of last week's midterm election, in which Republicans gained control of both houses of Congress, and with it the authority to lead all committees. Inhofe, who will be reclaiming a post he held from 2003 to 2007, is expected to use his chairmanship to attack President Barack Obama's climate and environmental policies, including his plans to reduce greenhouse gas emissions from coal-fired power plants and to strengthen protections for small streams and wetlands. The new Republican majority may also try to cut federal spending on social science research and open a high-level nuclear waste dump in Nevada—a project long stalled by outgoing Senate Majority Leader Harry Reid (D-NV).

In most policy areas of interest to the scientific community, however, the impact of the election is likely to be less obvious. The Senate's science and commerce committee, with a history of bipartisanship, will likely remain so when Senator John Thune (R-SD) takes over from retiring Senator Jay Rockefeller (D-WV). And when it comes to money, appropriations panel members from both parties have repeatedly pledged allegiance to increased funding for basic research, especially biomedical science.

The Republican takeover also isn't likely to splinter bipartisan coalitions that have emerged around a host of other research issues. They include efforts to renew a popular tax credit for business investment in research, cut research regulations and speed the development of new drugs, create a net-

work of advanced manufacturing centers, expand early education, and open the border to more high-skill immigrants.

Any efforts to boost funding will run up against mandated caps on overall federal spending, however (*Science*, 24 October, p. 408). And in spite of their majorities, Republicans have limited scope for unilateral action. The Senate's complex rules typically require 60 votes to pass legislation, and Republicans will have at most 54 (the race in Alaska is too close to call, and there's a runoff next month in Louisiana), preventing them from passing much of anything without Democratic help. And Obama's veto pen is poised. "Historically, divided governments tend to be more deadlocked than productive," says political scientist Sarah Binder of George Washington University in Washington, D.C.

Before Republicans can try flexing their new muscle in January, they must close out the year with a lame-duck legislative session in which Democrats still control the Senate. The top priority is to finish work on a 2015 budget for the fiscal year that began on 1 October. With spending now frozen at 2014 levels, science lobbyists are urging members to adopt a

draft bill that contains small increases for several research agencies, rather than continue the freeze until the new Congress—likely to be more tightfisted—convenes in January.

Budgetmakers in the House of Representatives and Senate also need to resolve a conflict over whether to withdraw U.S. support from the ITER international fusion project, now being built in Cadarache, France. A Senate panel proposed pulling the plug, but the House supports the project.

The current Congress may also take up the R&D tax credit, which expired at the end of 2013 for the sixth time in 21 years. The credit has funneled some \$8 billion annually to U.S. companies in recent years, and many law-

***"Historically, divided governments tend to be more deadlocked than productive."***

Sarah Binder, George Washington University political scientist



makers want to make it permanent. But that could cost \$100 billion to \$150 billion over the next decade, making a 2-year renewal more likely.

Once the Republicans take charge in the 114th Congress, social science researchers could be in for a tough ride, predicts John Porter, a former Republican representative from Illinois who championed funding for the National Institutes of Health (NIH) while chair of a House spending panel. "I think [social science funding] is dead at NIH and NSF [National Science Foundation]," he said at a postelection forum sponsored by Research!America, an advocacy group he chairs that's based in Alexandria, Virginia. "I just don't see how you're going to stop it from happening." Previous attempts to highlight what critics call wasteful or inappropriate government spending have won the support of some Democrats, whose votes would be needed again to impose such cuts.

Another veteran of the social-science wars, however, is less worried. "There's no need today to convince most people—in and out of Congress—that [understanding] behavior is important" to addressing national problems from obesity to terrorism, says Alan Kraut, head of the Association for Psychological Science in Washington, D.C. Although concerned, Kraut is confident researchers can head off any attacks.

One likely battleground is a bill, adopted last spring by the House science committee, that deals with NSF governance and the overall direction of federal research and education policy. The so-called FIRST (Frontiers in Innovation, Research, Science, and Technology) Act has so many provisions that scientists dislike—including changes to peer review and limits on funding for social sci-



President Obama hosted a lunch for congressional leaders last week.

are eager to revise rules that they say are stifling research, including reporting requirements for grants and approval processes for clinical trials.

One vehicle for change could be legislation from the bipartisan 21st Century Cures Initiative, backed by Representatives Fred Upton (R-MI) and Diana DeGette (D-CO), the top Republican and a senior Democrat, respectively, on the House energy and commerce panel. Still under wraps, the bill may call

for streamlining regulations at NIH and the Food and Drug Administration (FDA).

21st Century Cures has held dozens of hearings and roundtables, hearing testimony from an all-star list of biomedical research advocates, including NIH Director Francis Collins, senior FDA officials, and research lobbyists. The effort represents “the first time in quite a while where the [scientific] community has been asked to try and identify specific things that they think Congress can do,” says David Moore, senior

director of government relations at the Association of American Medical Colleges in Washington, D.C.

Another potential bright spot for the research community is advanced manufacturing, an issue on which Obama and the Republican majority largely agree. This fall, the House passed a bill that closely mirrors the president’s own position on creating a network of research centers that would also promote regional economic development. The one area of disagreement is the funding mechanism—Republicans would take the money from an existing energy efficiency program at the Energy Department instead of tapping new revenue from the sale of public communications spectrum. A similar, bipartisan bill is awaiting Senate action, and reviving it in the next session could lead to an easy victory for both parties.

As Washington awaits the new Congress, Porter had some advice for federally funded researchers hoping to stay out of any political crossfire: Be careful how you write abstracts. He was alluding to recent episodes in which Republican lawmakers have publicly criticized and ridiculed some NIH grants for failing to explain how they might improve human health. “You’d think the scientists, who are pretty smart, would figure it out,” Porter says, suggesting that NIH Director Collins “get some people to look over those things before they are made public.” ■

*With reporting by the Science News staff.*

ence—that many science lobbyists hope Congress will simply shelve it unless Thune can lead his Senate colleagues to some middle ground. Help forging such a compromise could come from the expected new chair of the Senate’s health, education, and labor committee, Senator Lamar Alexander (R-TN).

In contrast, parts of the science community are likely to welcome another item on the Republican agenda: cutting government regulation. In particular, universities

## Will pork be back on the menu?

By Jeffrey Mervis

Last week’s election could herald the return of earmarks—Congress’s controversial practice of funding projects proposed by individual legislators that are not in the president’s budget request.

Members traditionally used earmarks to build roads and bridges in their states. But more recently earmarks also financed research projects and facilities—a practice that most scientists believe subverted the merit-review process. Republicans in the House of Representatives banned them after reclaiming the majority in 2010, and the Democratic Senate grudgingly went along. But with voters angry at Congress for the current gridlock, some party leaders favor bringing back earmarks, or “pork,” to help grease the lawmaking process.

There’s little question that pork can influence votes. In 1964, for instance, “we wouldn’t have had a Civil Rights Act if President [Lyndon B.] Johnson hadn’t promised Purdue [University] a NASA research center,” says Jason Grumet, president of the Bipartisan Policy Center, a good-government think tank in Washington, D.C. The NASA earmarks (actually a series

of research grants to faculty) won over Representative Charles Halleck, the Republican House leader whose Indiana district included Purdue. Earmarks are part of what Grumet calls “constructive partisanship,” and last week two former Senate majority leaders endorsed the practice during a postelection roundtable hosted by Grumet’s center.

The absence of earmarks “has made it harder to pass legislation,” says former Senator Trent Lott (R-MS). “I’ve always been in favor of them, and I know Senator [Richard] Shelby would also like to get them back,” he adds, referring to the Republicans’ top appropriator in the Senate. Senator Mitch McConnell (R-KY), the presumptive Senate majority leader in the next Congress, is another notorious earmarker.

Former Senator Tom Daschle (D-SD), who swapped jobs with Lott for nearly a decade as control of the Senate kept flipping, says bringing back earmarks would also allow legislators to reclaim their constitutional power over federal spending: “I’d much rather have an elected official make these [spending] decisions than an unelected one,” he says, referring to agency appointees.

Lott isn’t sure the new Senate leaders will take the political risk of restoring the practice. If it is revived, both Lott and Daschle say, it must be accompanied by a more “transparent” process. ■



## SCIENTIFIC COMMUNITY

# No sexism in science? Not so fast, critics say

## Study of hiring in math-heavy fields provokes outcry

By Rachel Bernstein

There has never been a better time for American women to enter academic careers in math-intensive science fields. That's the message Cornell University psychologist Stephen Ceci says he was hoping to get across in an op-ed that he and Wendy Williams, also a Cornell psychologist, recently published in *The New York Times*. But the 31 October article, based on a new study of U.S. government data and existing literature co-authored with economists Donna Ginther of the University of Kansas, Lawrence, and Shulamit Kahn of Boston University, left some readers unconvinced and others outright angry.

The op-ed ran under a provocative headline: "Academic Science Isn't Sexist." It asserted that "the experiences of young and midcareer women in math-intensive fields are, for the most part, similar to those of their male counterparts" and that female underrepresentation in some fields is "rooted in women's earlier educational choices, and in women's occupational and lifestyle preferences." The outcry on blogs and Twitter was swift. But some accepted the authors' findings about how women are faring in physics, math, engineering, and similar fields, while rejecting their optimistic analysis. "It's not that I don't think their data is accurate," says sociologist Jennifer Glass of the University of Texas, Austin. "It's that I think their interpretations should be taken with a grain of salt."

Most agree that, in U.S. universities today, things are much better for women than they used to be. In their study, published in the current issue of *Psychological Science*

in *the Public Interest*, the researchers tried to quantify the improvement by reviewing relevant data on academic hiring and promotion, mainly from the biennial Survey of Doctorate Recipients conducted by the National Science Foundation. They found that women who compete for assistant professor positions in math-intensive science fields are just as likely to be hired as men are, if not more so, and that women are not discriminated against in tenure and promotion decisions. "Although women are underrepresented in math-based fields, those women who go into them do as well as the men who go into them," Ceci says.

"That's really something to celebrate," says psychologist Diane Halpern of the Keck Graduate Institute in Claremont, California, who wrote a commentary that accompanies the paper. Others, though, are more cautious. "The problem with observational data is that you can't determine cause and effect very easily," says psychologist Virginia Valian of the City University of New York's Hunter College. As Glass puts it, "Measures of equal performance or equal opportunity in hiring do not mean there is no bias. They mean that women have overcome any bias that may exist."

Corinne Moss-Racusin has evidence that bias is still present. As a postdoc at Yale Uni-

versity, the psychologist was the lead author of a study published in the *Proceedings of the National Academy of Sciences* in 2012 showing that both male and female science professors evaluate a resume more favorably if they believe that the applicant is male. "We're able to conclude pretty convincingly that bias exists," says Moss-Racusin, who is now at Skidmore College in Saratoga Springs, New York.

Ginther acknowledges that "it's very difficult to uncover bias in the [correlational] data" that were the basis of the current study. But "we don't see a lot of evidence [for it]," she says. To explain the striking gender imbalance in the studied fields—as of 2010 women

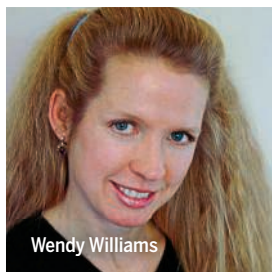
composed only 7% to 16% of full professors in math-intensive fields—she and her colleagues point instead to women's earlier choices about, say, which high school math classes to take and what college major to enroll in. But some other researchers complain that the authors ignore the cultural pressures that influence such choices. "These are not free choices," says psychologist Janet Hyde of the University of Wisconsin, Madison. "Sure, women make choices, but the choices are heavily constrained by environmental factors."

Ginther notes that she and her co-authors did not

intend to blame women for their choices. Instead, they hope to identify the point at which intervention could be most effective. "If you're going to treat a symptom"—the underrepresentation of women in math-intensive fields—"you've got to know where it's happening and focus your efforts on where things start to diverge." One of the paper's main recommendations is for math interventions for girls at the precollege level.

Ceci also hopes that the results will encourage young female scientists by counteracting the negative stories they are likely to hear about science. "It would be very unfortunate if a talented young woman opted out of a career in science because she read anecdotal reports or testimonials that say it's a hostile work environment."

Others, though, worry that the paper and the accompanying op-ed could hurt the ongoing effort to increase gender balance in the sciences. Hyde is concerned that downplaying the challenges young women are likely to face leaves them unprepared for the real environment in academia. "They're going to hit their first instance of sex bias and they're going to be totally floored." ■



Wendy Williams

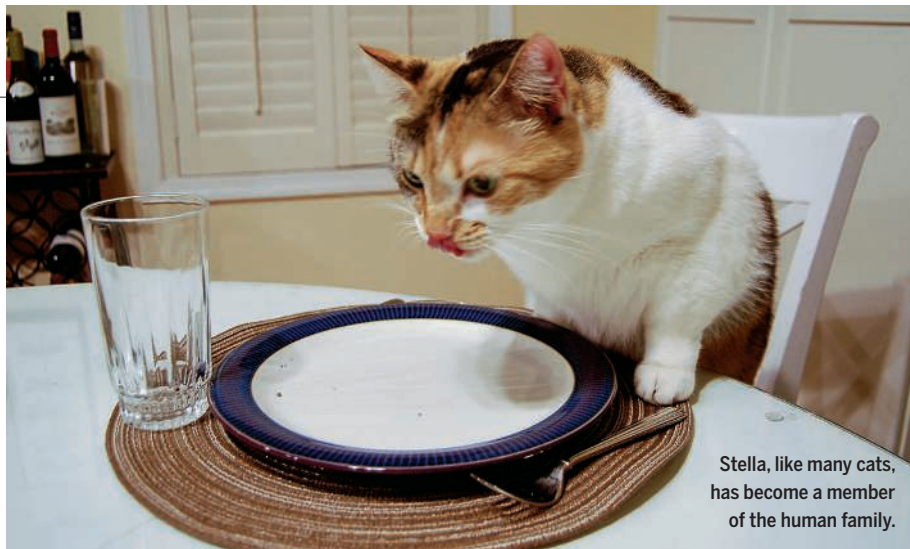


Stephen Ceci

### A level playing field for women?

At 89 U.S. universities, women were interviewed and hired for tenure-track jobs at rates matching or exceeding the proportion of female applicants.

| FIELD       | PERCENT OF APPLICANTS | PERCENT OF CANDIDATES INTERVIEWED | PERCENT OF CANDIDATES OFFERED POSITION |
|-------------|-----------------------|-----------------------------------|--|
| Physics     | 12%                   | 19%                               | 20%                                    |
| Biology     | 26%                   | 28%                               | 34%                                    |
| Chemistry   | 18%                   | 25%                               | 29%                                    |
| Mathematics | 20%                   | 28%                               | 32%                                    |



Stella, like many cats, has become a member of the human family.

## ANIMAL DOMESTICATION

# The genes that turned wildcats into kitty cats

Findings help show how all animals became tame

By David Grimm

Place a housecat next to its direct ancestor, the Near Eastern wildcat, and it may take you a minute to spot the difference. They're about the same size and shape, and, well, they both look like cats. But the wildcat is fierce and feral, whereas the housecat, thanks to nearly 10,000 years of domestication, is tame and adaptable enough to have become the world's most popular pet. Now scientists have begun to pinpoint the genetic changes that drove this remarkable transformation. The findings, based on the first high-quality sequence of the cat genome, could shed light on how other creatures, even humans, become tame.

"This is the closest thing to a smoking gun we've ever had," says Greger Larson, an evolutionary biologist at the University of Oxford in the United Kingdom who has studied the domestication of pigs, dogs, and other animals. "We're much closer to understanding the nitty-gritty of domestication than we were a decade ago."

Cats first entered human society about 9500 years ago, not long after people first took up farming in the Middle East. Drawn to rodents that had invaded grain stores, wildcats slunk out of the deserts and into villages. There, many scientists suspect, they mostly domesticated themselves, with the friendliest ones able to take advantage of human table scraps and protection. Over thousands of years, cats shrank slightly in size, acquired a panoply of coat colors and patterns, and (largely) shed the antisocial tendencies of their past. Domestic animals

from cows to dogs have undergone similar transformations, yet scientists know relatively little about the genes involved.

Researchers led by Michael Montague, a postdoc at the Washington University School of Medicine in St. Louis, have now pinpointed some of them. The scientists started with the genome of a domestic cat—a female Abyssinian—that had been published in draft form in 2007, then filled in missing sequences and identified genes. They compared the resulting genome with those of cows, tigers, dogs, and humans.

The analysis, published online this week in the *Proceedings of the National Academy of Sciences*, revealed 281 genes that show signs of rapid or numerous genetic changes—a hallmark of recent selection—in domestic cats. Some appear to be involved in hearing and vision, the senses that felines

rely on most. Others play a role in fat metabolism and are likely an adaptation to cats' highly carnivorous lifestyle.

But the most intriguing findings came when the team sequenced the genomes of 22 domestic cats—representing a wide variety of breeds and locations—and compared them with the genomes of two Near Eastern and two European wildcats. The researchers uncovered at least 13 genes that changed as cats morphed from feral to friendly. Some of these, based on previous studies of knockout mice, seem to play a role in cognition, including fear responses and the ability to learn new behaviors when given food rewards. "That jibes with what we know about the domestication of cats," Montague says, "because they would have needed to become less fearful of new locations and individuals, and the promise of food would have kept them sticking around."

"This is my favorite part of the paper," says Kerstin Lindblad-Toh, a leading comparative genomicist at Uppsala University in Sweden who was not involved in the work. She notes that a few of the genes the team identified code for glutamate receptors, which play a key role in learning and memory and may have been selected in humans as well. "We're hitting on genes that allow our brains to develop and make us interact socially."

The team also found five genes in domestic cats that influence the migration of neural crest cells, stem cells in the developing embryo that affect everything from skull shape to coat color. This supports a recent proposal that such cells may be the master control switches of domestication, explaining why domestic animals share common traits, such as smaller brains and certain pigmentation patterns—a mystery first noted by Charles Darwin (*Science*, 24 October, p. 405).

So why are cats still a bit wilder than our other favorite domesticate, the dog? Co-author William Murphy, a geneticist at Texas A&M University, College Station, says the cat genome appears to have undergone less intense and more recent evolutionary pressure than that of dogs; that's not surprising, considering that dogs may have lived with us for up to 30,000 years. "Cats were not selected for a purpose like dogs and other domesticates," Murphy speculates. "They just hung out, and humans tolerated them."

Still, Larson doesn't think it's fair to call cats "semi-domesticated," as the authors do in their paper. "I've got two cats at home, and they're as domesticated as any animal on Earth," he says. "There are homes where cats just sit on the couch, ignoring the dogs and primates that should be a major threat to them. That's asking a lot of a wild carnivore." ■





## SCIENCE AND CULTURE

# The theoretical physicist behind *Interstellar*

What Kip Thorne loved about the new science-fiction movie—and what made him cringe

By Daniel Clery

*Interstellar*, which opened last week, boasts not just a compelling storyline and dazzling special effects but also an impeccable scientific pedigree. The man who inspired the film and kept a close eye on its scientific fidelity is Kip Thorne, 74, a renowned theoretical physicist at the California Institute of Technology in Pasadena and one of the world's leading experts on the astrophysical predictions of general relativity.

In 2006, Thorne and Lynda Obst, a longtime friend and film producer, wrote an eight-page treatment for a film that sprang from the astrophysics of black holes, wormholes, and time dilation. Steven Spielberg was soon on board to direct. Six years later, however, Spielberg had to drop out and was replaced by Christopher Nolan, director of the three *Dark Knight* movies and *Inception*. His brother Jonathan Nolan, who wrote films including *The Dark Knight Rises*, worked on the screenplay, which is set in a not-too-distant future, when crop

blights have driven humanity to the brink of starvation. A secret, crash attempt to find another planet that humans could colonize is under way.

Thorne writes about his experiences working with Hollywood and the scientific concepts addressed in the film in a book, *The Science of Interstellar*, published on 7 November. He spoke with *Science* last week. This interview has been edited for clarity and brevity.

**Q: How much of your original treatment remains in the final film?**

**A:** It's not recognizable as the same movie except for the scientific vision and the venue, what I like to call the warped side of the universe—black holes, wormholes, higher dimensions, and so forth.

The story is essentially completely changed, except in the broadest brush sense that we had explorers leaving the Earth, traveling out to a wormhole in the outer parts of the solar system and through the wormhole, and visiting planets. The thing that was preserved was the vision—the thing that was most important to Lynda and me—the vision of a movie in which real science, ranging from truth to speculation, is embedded deep into the movie's fabric from the outset.

Also preserved were the guidelines we laid down. First, that nothing would violate established physical laws. Second, that all the wild speculations, and there certainly are some here, would spring from science and not from the fertile mind of a screenwriter. I'm very pleased with how it came out.

**Q: Is there anywhere the moviemakers strayed outside your guidelines?**

**A:** Not seriously. The one place where I am the least comfortable is on [a] planet where they have these ice clouds. These structures go beyond what I think the material strength of ice would be able to support. Every time I watch the movie, that's the one place where I cringe. I don't think I've ever told anybody that. But I'd say if that's the most egregious violation of physical law, they've done very, very well.

**Q: In your book, you say that Christopher Nolan introduced science elements of his own to the script. What did he bring?**

**A:** The one that has the biggest impact, and that I really like, is the tesseract [the 4D analog of a cube]. When he told me in our first meeting that he was thinking of using a tesseract, he didn't go into any detail. But I was very pleased because when I was 13 years old I read a book, *One Two Three ... Infinity* by George Gamow, in



Theoretical physicist Kip Thorne arrives at the Hollywood premiere of *Interstellar* last month.

which George drew a picture of a tesseract. It looks like two cubes, one inside another, and I spent hours staring at that, trying to come to grips with it. I found it so fascinating that it was one of the more significant influences on me to become a theoretical physicist.

So when [Nolan] told me he wanted to use a tesseract, I immediately saw and discussed with him how this was an ideal way to take his hero and carry him into the fifth dimension, and rapidly travel from one region of our universe to another, because distances in the [fifth dimension] will be much less than they are in our brane [our 4D slice of reality].

**Q: Where did the idea of blights plaguing Earth come from?**

**A:** This was due to [Jonathan Nolan]. When he introduced the idea, he and I and Lynda decided we really needed to find what was known about blight and other kinds of biological catastrophes. We set up a “blight dinner” with biologists who were experts and we discussed this at length, trying to identify what things could go wrong with the biology of the Earth.

**Q: You worked with the visual effects team at the company Double Negative in London, providing them with equations, which they then worked into their code. What was it like seeing those equations turned into the visual representation of a black hole?**

**A:** It was wonderful to see the resolution they got. I knew roughly what they were going to come up with, but it was just awe-inspiring to get back film clips from [Double Negative] and see this fabulous resolution and fabulous dynamics they were able to achieve.

**Q: You’ve said that you learned something new from their simulations?**

**A:** We learned [that] when you have a fast-spinning black hole, without any accretion disk, and let it just lens the distant sky—a star field—we saw a fantastically beautiful structure that is sort of like a fingerprint, but much more complex. It was completely unexpected with huge amounts of internal structure in it, regions where the star field appears to be quiescent and other regions where the stars seem to be whirling around in little vortices. To me it’s a lovely kind of discovery in the sense that it is really very beautiful and it arises from a collaboration between a scientist and a group of computer artists. We are submitting a paper about this and about the particular method that Double Negative uses to the journal *Classical and Quantum Gravity*. ■

## MICROBIOLOGY

# Contamination plagues some microbiome studies

Extraneous DNA can skew results of surveys of low-density microbial communities

By Elizabeth Pennisi

**F**ive years ago, Susannah Salter and her colleagues swabbed the throats of 20 infants monthly until the children reached 2 years old—and documented a remarkable change. By sequencing the swabs’ DNA, Salter, a microbiologist at the Wellcome Trust Sanger Institute in Hinxton, U.K., learned that infants and toddlers have quite different microbial communities in their throats. Or so it seemed, until the team discovered that change was an artifact of contamination in two different batches of the DNA isolation kit used in the study. “The pattern wasn’t real,” says Salter’s former Sanger colleague Alan Walker, now a microbiologist at the University of Aberdeen in the United Kingdom.

Worried that many microbiome researchers are falling into the same pitfall, Salter, Walker, and their colleagues have now done a systematic study of how contamination can affect DNA survey results. They find that most testing kits and other reagents for analyzing DNA contain stray microbial DNA that can overpower the signal of the native bacteria. Their study, reported this week in *BMC Biology*, “is of huge significance, because it either implicitly or explicitly debunks quite a large number of microbiome papers,” says Mick Watson, a bioinformaticist at the Roslin Institute of the University of Edinburgh in the United Kingdom. “I think everyone kind of knew that there might be contamination in reagents and kits, but no one really knew the extent of the problem.”

To perform their test, Salter, Walker, and their colleagues cultured a single *Salmonella* strain and then dissolved the bacteria in an otherwise sterile broth. Then

they repeatedly diluted the broth, testing it each time using several DNA isolation kits and other laboratory reagents and methods. Early in the process, when the sample bacteria were abundant, the small proportion of stray DNA in the test kits was not enough to skew the outcome, but when the sample was diluted to just 1000 bacterial cells, the contaminant DNA swamped the real signal, the team reports.

The team thinks such contamination will not be a problem for microbiome studies of fecal samples, as they are teeming

with gut bacterial DNA. But it looms large for samples typically containing relatively few bacteria, such as dust, spinal fluid, or blood, Walker says.

Manufacturers of DNA isolation kits don’t guarantee their materials are sterile, he notes, but many people might not realize that. “This study will have a big impact on how we design and carry out our lab protocols and sequencing studies in the future,” says microbial ecologist Holly

Bik of the University of Birmingham in the United Kingdom.

Veterans of microbiology have for decades warned about the need to take contamination precautions seriously and exchanged horror stories. In 1998, for example, scientists were finding the same microbes in all sorts of extreme environments; it turned out the species in common came from water pipes. This “issue is as old as the field itself and it’s being continuously rediscovered,” says Ruth Ley, a microbial ecologist at Cornell University. And it can be dealt with, says Gary Huffnagle, who studies lung microbes at the University of Michigan, Ann Arbor. “This study should not be used to dismiss analysis of low biomass samples,” he says. “Running controls is always key.” ■



Contamination may spoil surveys of microbes in throats and other tissues.





# The RIVER MASTERS

Hippos are the nutrient kingpins of Africa's waterways

By **Elizabeth Pennisi**, at the Mpala Research Centre in Kenya

In the cool, early morning light, the hippos are at peace. Just the tops of their massive heads are above the water, mist streaming from their nostrils as they exhale. One rests its chin on a rock. Another wiggles its ears and then gently submerges, leaving a circle of expanding ripples. A third yawns, exposing four large canines, muscular lips, and a giant, bright pink mouth.

To ecologist Douglas McCauley, this scene on the Ewaso Ng'iro river in Kenya is far

more than a safari postcard. Unseen, massive amounts of nutrients are entering the river, helping sustain the ecosystem: The hippos are pooping.

For the complex food webs of sub-Saharan Africa, hippo dung is an elixir. Just as salmon swim upstream in northern latitudes to spawn and die, infusing rivers with the stuff of life from their own decaying bodies, hippo dung transfers energy from land, where the animals graze, to Africa's turbid rivers, fueling growth of the organisms at the

base of the food web. "We think of hippos as megaton salmon," says McCauley, of the University of California (UC), Santa Barbara. By measuring telltale ratios of carbon isotopes, McCauley, working with Justin Brashares, a conservation biologist at UC Berkeley, and others, has traced the flow of energy through the food web, from hippo dung on up to crayfish and barbel fish. Their studies have revealed that hippos transport so much carbon and other nutrients in their prodigious excrement that they can support—or poison—an entire aquatic ecosystem.

Researchers once shunned hippos in the wild, deeming the scientific payoff too small for the mortal risk of getting up close with the dangerous beasts. The extraordinary story of hippo poop is changing minds. "Up until now there's been a great deal of speculation and back-of-the-envelope calculations," says Robert Naiman, an ecologist emeritus at the University of Washington, Seattle. "Finally we have quantitative evaluations of the role of hippos," he says—and a case study of the powerful ecological role a single large animal can play.

PHOTO: © ANGELO CAVALLI/CORBIS

Downloaded from www.sciencemag.org on November 17, 2014





**MCCAULEY'S TOUSLED SANDY HAIR** contrasts sharply with the crisp white dress shirt he's wearing to avoid sunburn on the sun-drenched savanna. When he first encountered a hippo in the wild in 2002, he recalls, "I was scared to death." Ornerly and territorial, these aquatic herbivores are considered by many to be Africa's deadliest animal, reputedly killing more people than elephants or lions. Weighing up to 4 tons, they can sprint as fast as a person, crushing victims to death with their powerful jaws or by trampling them. And although McCauley used to think he was safe on land when hippos are in the water, one day when he was sampling fish from the riverbank a hippo surfaced ever closer and then suddenly charged up the bank, forcing him to make a dash for his Land Cruiser.

Apart from their reputation for inflicting painful forms of death, relatively little is known about how hippos interact with each other, with other creatures, and with the environment. "It seems like at this point we shouldn't be at this square-oneish level, but with hippos we really still are," complains

Simply by eating and excreting, the hippopotamus can transform entire African ecosystems.

Rebecca Lewison, a conservation biologist at San Diego State University in California and chair of the International Union for Conservation of Nature Hippo Specialist Group. "It's pathetic."

One reason for the dearth of knowledge is the acute challenge of telling one hippo from another. This close relative of cetaceans, McCauley says, is like "a cross between a baleen whale and a pig"—two well-studied species. But whereas biologists can distinguish individual whales by fin and tail markings, they have struggled to work out whether hippo ears, for example, have unique outlines, or whether individuals can be identified from scarring patterns, face color, or whisker placement.

Merely observing the beasts is a challenge because they pass their days submerged in lakes, ponds, and eddies in muddy rivers. At dusk, they "swim" by bouncing off the river bottom and eventually clamber ashore. They graze in grassy "glades" at night—when snakes, elephants, lions, and hyenas make it unwise for a researcher to be out—and slip back into the water just before dawn. Nor can researchers easily fit them with the GPS collars used to track many kinds of African mammals for behavioral studies. Most collars slip right off a hippo, which doesn't really have a neck.

Last year, Tristan Nuñez, a graduate student at UC Berkeley, came up with an alternative strategy for tracking hippos. He took aim at basking hippos with his crossbow—but instead of arrows he fired barbed GPS tags of the sort used to track whales. It was hard to get a clean shot, he says, but twice he stuck a tag in a hippo's tough hide. Be-

ing able to track an individual hippo's movements was "a major triumph," Lewison says. But for now, McCauley says, "we don't have a sense of their patterns of movements," such as how far they wander at night for food.

To get answers, McCauley more than once climbed a tree overhanging a hippo pool and spent the day taking notes. And in 2012, he and Mordecai Ogada, now with the Laikipia Wildlife Forum in Nanyuki, Kenya, and their colleagues used camera traps to collect

50,000 hours of footage of hippo grazing areas. The team also drove 9 hours on a treacherous highway to Kenya's Mzima Springs, where clear water allowed them to observe hippos underwater both from the bank and from a partially submerged observation

room for tourists. "They are quite comical ... and graceful for being so big," says McCauley, who gained a new appreciation for the hippo ballerinas in the movie *Fantasia*. At Mzima, they watched young hippos romp like 2-year-old children and adult males establish dominance by spraying dung in the face of rivals.

**TO PROBE THE ROLE** hippos play in the savanna ecosystem, McCauley and his colleagues also established five sets of 30-square-meter plots, monitored by cameras, along the Ewaso Ng'iro river. Some plots are edged with evenly spaced poles that are just the right size to keep out hippos while allowing other animals through. Electrified wiring around other plots excludes most animals, and some plots have no fence at all. Twice a year, McCauley's team surveys plants in the plots and clips leaves and scoops soil to analyze nutrient content.

The experiment has revealed that hippos reshape the grasslands to their advantage. In plots they frequent, grass is stubbier and leafier because of repeated cropping. That's good for hippos, McCauley says: "Getting more leaf area per bite is better because it's much more nutritious," he says. In hippo-free plots, grass is more strawlike and flowers more frequently.

Given that a single hippo eats about 40 kilograms of grass a night, McCauley wanted to know how the animals' daily perambulations transport nutrients such as carbon, phosphorus, and nitrogen through the environment. Past studies of nutrient transport in ecosystems have focused on physical forces such as runoff and on the influence of plants and microbes. "Animal

#### PODCAST

To hear a podcast with author Elizabeth Pennisi, see [http://scim.ag/pod\\_6211](http://scim.ag/pod_6211).



Douglas McCauley sets up a camera trap to snap photos of hippos as they graze at night.



behavior has been left on the sidelines,” says David Post, a community ecologist at Yale University. “Yet animals can move [nutrients] faster than physical forces or even in a way that’s counter to these forces,” he says.

To quantify the nutrient flows in African ecosystems, McCauley and his colleagues turned to isotopic analyses. Different plant species have characteristic proportions of heavy and light carbon isotopes in their tissues. Each plant’s isotopic ratio provides a marker for its carbon as it works its way through the food web. Researchers can also add tracer isotopes to food for more precisely measuring the energy transfer from organism to organism.

Tracer isotopic analysis has unraveled nutrient use in birds, pigs, and sheep, but its use in wild animals would be unprecedented, says Todd Dawson, a plant ecologist at UC Berkeley, who led the isotopic analyses for McCauley’s team. First they tried out the technique with captive hippos at Disney World in Florida. They fed the hippos food with known isotopic compositions and tracers and confirmed that the excrement preserved the same signatures. Once they proved they could track energy flow in this way, they fed hippo dung to guppies. Unpublished results show that the guppies’ isotopic ratios mirrored those of the dung.

Back in Kenya, the team determined isotopic ratios of local plants. Then they went hunting for hippo dung. “It’s like an Easter egg hunt,” McCauley says. “There’s not much there, but that’s the point—not a lot of dung winds up on the land.”

What’s more, hippos don’t just poop. They swish their tails as they go, sending dung flying into the bushes. On one outing, McCauley visited two sites and spent 45 minutes before finding what he was looking for: fresh dung, still warm, splattered in the sand. He carefully wrapped the prize in aluminum foil to carry back to the Mpala Research Centre, an institute near Mount Kenya that serves as McCauley’s home and workplace while in Kenya.

From isotopes in dung samples, his team has learned which plant species hippos like to eat, and how hippo grazing skews the species assemblage of a glade. Next they hope to track how eating habits vary between wet and dry years and between

seasons, which would indicate how hippos influence land eco-systems over time.

They are also using the isotopes to follow the nutrient chain further: from hippo poop to the fish and insects it nurtures. Already they have discovered that its nutritional impact on an ecosystem “is very strongly regulated by what the river is doing,” McCauley says. Where currents are swift and hippos are few, the dung makes little difference to the river’s other inhabitants. But come dry season, lower water flow leads to pooling, and isotope studies of river creatures show a clear signature of hippo poop. McCauley says that even as few as a dozen hippos “will become a dominant contributor to the nutrient pool.” The fish they tested are an important commercial species in some Kenyan lakes, he notes, suggesting that hippos “need to be recognized as a life force for the industry.”

But researchers have also discovered that hippo dung can be too much of a good thing.

**ALONG THE MARA RIVER**, a key waterway that meanders through the Serengeti, some 4000 hippos are crowded into 155 kilo-

meters of river. The throng has increased 15-fold in the past several decades, likely in part because hunting is now illegal. At the same time, the river’s water quality has deteriorated. When Amanda Subalusky and her husband, Christopher Dutton, came to Africa in 2008 to investigate, they flagged deforestation as a major water despoiler. But to their surprise, water quality was poorest in protected areas, particularly when flows were low. The common factor appeared to be hippo habitat.

To probe further, Subalusky, now a Yale graduate student, and Dutton, who studies with Emma Rosi-Marshall at the Cary Institute of Ecosystem Studies in Millbrook, New York, paid a visit to the Milwaukee County Zoo, home to three hippos. Zookeepers drain the hippo pool daily, allowing Dutton and Subalusky to quantify the urine and feces produced each day and determine the amount of nitrogen, phosphorus, and carbon the hippos egested. Extrapolating to the Mara, they will report in an upcoming issue of *Freshwater Biology* that 4000 hippos would contribute more than 36 tons of feces a day.

The Mara can’t always cope with such abundance. Subalusky and her colleagues have linked sudden pulses of water after heavy rainfalls with precipitous declines in dissolved oxygen and fish kills, most likely brought on when the current stirs up thick beds of decomposing dung on the river bottom. Her team will test that idea with a boat that was built by Platypus LLC, a spinoff from a Carnegie Mellon University robotics group. The vessel looks like a crocodile (on the hunch that hippos will steer clear), and it should allow them to measure pool depth and dung deposition before and after a flood. “It’s a chance to see where ecosystem dynamics are really influenced by an animal,” Subalusky says.

McCauley witnessed a similar phenomenon in southern Tanzania’s Ruaha River Basin. Until the late 1990s, its waters flowed year-round. But increased irrigation demands for rice have reduced the flow, and the river now breaks up into pools much of the year. Some pools teem with 100 or so hippos, while others have none. In the hippo-clogged pools, nutrients from dung can set off an algal bloom, and as the

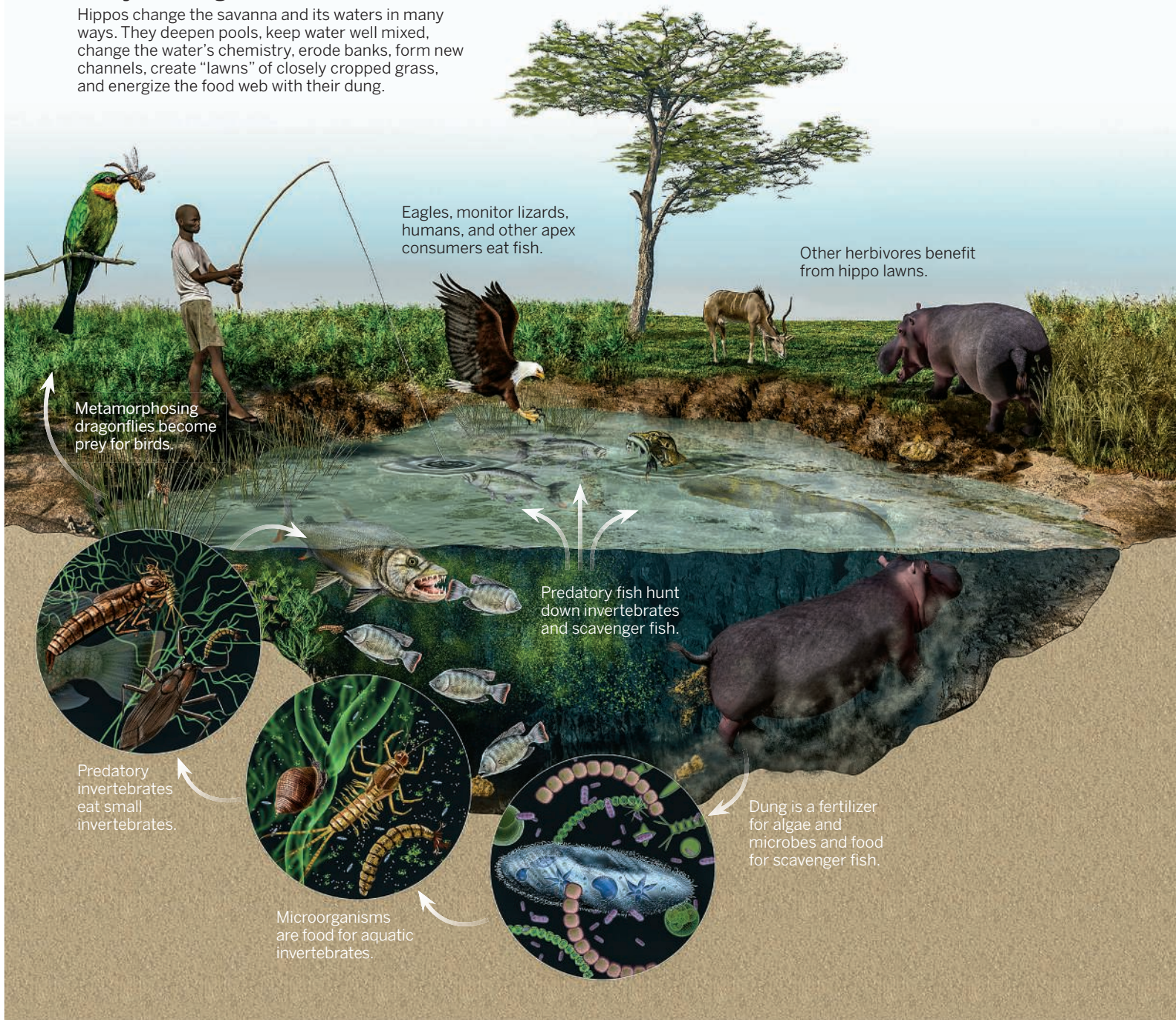


A Milwaukee zoo pool that went from clean to filthy in less than a day allowed Amanda Subalusky to measure nutrient content of hippo dung and urine.



## Ecosystem engineers

Hippos change the savanna and its waters in many ways. They deepen pools, keep water well mixed, change the water's chemistry, erode banks, form new channels, create "lawns" of closely cropped grass, and energize the food web with their dung.



dying algae sink and decompose, they draw down the oxygen, turning the water black. "There are few river animals that can cope with such conditions," McCauley says.

Despite their abundance in the Mara and Ruaha rivers, hippos broadly are on the decline. About 100,000 years ago, they roamed as far north as the United Kingdom and across much of Europe. They disappeared from Egypt a century ago, and now they are confined to sub-Saharan Africa. Even there

they are threatened. In Mozambique, for example, relentless hunting and losses during the long civil war that ended there in 1992 reduced a population numbering in the thousands to just a few hundred. Today, an estimated 135,000 remain in all of Africa. Unlike elephants, which are also in steep decline, hippos "do not have a champion," says San Diego State University's Lewison.

The new portrait of the hippo as the unsung nutrient kingpin of sub-Saharan

Africa suggests that as the animals decline, ecosystems will be transformed, affecting many other creatures. That prospect adds urgency to the efforts of the handful of scientists who are striving to paint a more nuanced picture of the formidable mammal. "Scientifically, we've overcome quite a few challenges," says UC Berkeley's Dawson. The next challenge, he says, is making the science relevant to conservation. ■





## PERSPECTIVES

### ENVIRONMENTAL SCIENCE

# The trouble with neonicotinoids

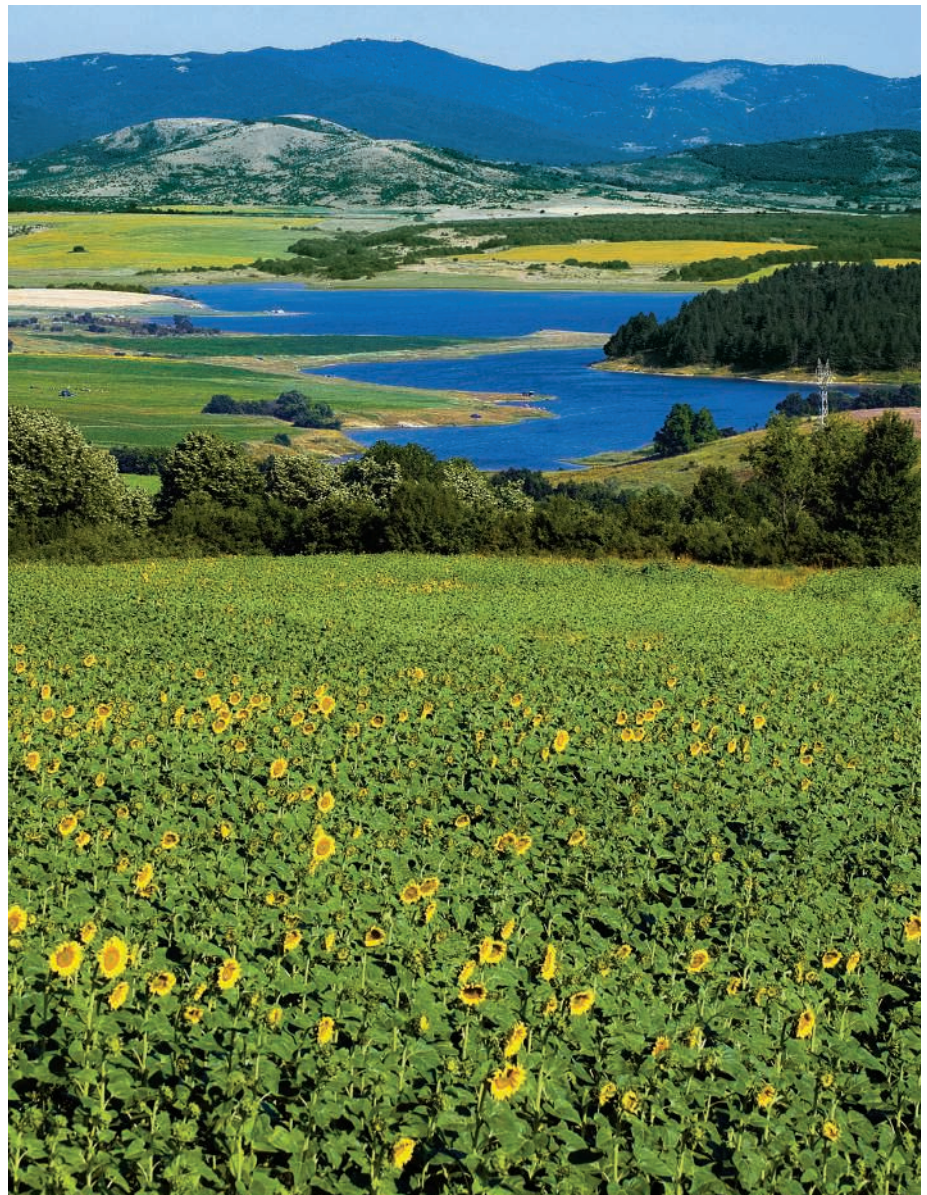
Chronic exposure to widely used insecticides kills bees and many other invertebrates

By **Francisco Sánchez-Bayo**

**F**our decades ago, DDT and other pesticides that cause environmental harm were banned. Since then, newly developed pesticides have had to conform to stricter environmental standards. Yet, recent studies highlight the subtle but deadly impacts of neonicotinoids—the most widely used insecticides in the world—on ecosystems (1–3). In contrast to other insecticides, neonicotinoids are systemic, meaning that they are highly soluble and thus absorbed by the plant. They produce delayed mortality in arthropods after chronic exposure to sublethal doses but are not very toxic to vertebrates. It has taken more than a decade to unravel some of the mechanisms through which neonicotinoids affect the integrity of ecosystems. Although gaps in knowledge remain, there is a strong case for stricter regulation of these pesticides.

Neonicotinoids are mainly applied as granules into the soil or as seed-dressings during crop planting. Seeds are coated with 1 to 17 mg per kg, depending on crops and compounds. As plants grow, they take up 2 to 20% of the insecticide and distribute it to all parts of the plant, including leaves, flowers, pollen, and nectar. The resulting concentrations of 5 to 10 µg per liter [parts per billion (ppb)] in the sap are sufficient to control sucking and chewing insect pests (see the figure). However, pollinators such as bees, butterflies, moths, and hoverflies are equally exposed; where neonicotinoids are used, 11 to 24% of pollen and 17 to 65% of nectar is contaminated with these insecticides (3).

Soon after the neonicotinoid imidacloprid was introduced in France in 1994, beekeepers noticed that their honey bee colonies were weakening or disappearing. The ensuing investigation found that this and another systemic insecticide (fipronil) were



**Hidden killers.** Neonicotinoids applied to seeds or soils spread into the environment, killing many nontarget arthropods.

PHOTO: MARA RADEVA/THINKSTOCK

particularly toxic to bees, with acute dietary LD<sub>50</sub>'s (dose to kill 50% of bees) of 2.5 to 5 ng per bee (4). Forager bees do not die immediately after visiting flowers in treated crops because residue levels are below their acute LD<sub>50</sub> and bees only ingest part of what they collect; the rest is taken to the hive. It is the daily sublethal doses the pollinators ingest that are the problem. Effects include olfactory learning, memory, and locomotory impairment and inhibited feeding (5). In a laboratory study, chronic ingestion of 4 to 8 ppb imidacloprid resulted in 50% survival of honey bee workers after 30 days (6).

Whether these observations apply to bees in the natural environment has been a contentious question, because the performance of the hives does not change significantly. There are several reasons for this apparent

weakens the bees' immune system, making them more susceptible to pathogens such as *Nosema*. These confounding factors can be blamed for the declines in honey bees but cannot account for the parallel decline in wild and bumble bees.

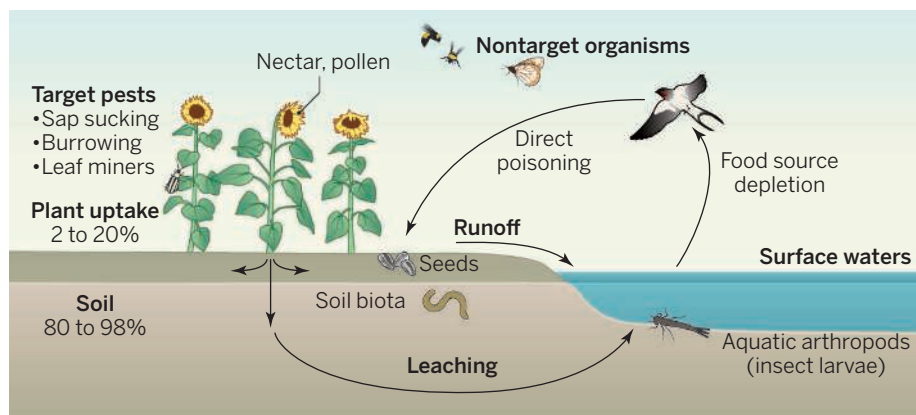
Di Prisco *et al.* (1) have established that sublethal doses of two neonicotinoids (clothianidin and imidacloprid) cause bee immune deficiency that triggers viral infections. This causal link helps to explain the time lag between initial exposure and mortality: The cascade of effects prompted by the insecticides involves irreversible biological pathways that are not observable until death takes place (9).

Although bees have captured most of the attention, neonicotinoids are equally toxic to ants, termites, parasitoids, and aquatic

Feeding inhibition has been observed in several decomposer organisms exposed to chronic, sublethal concentrations of imidacloprid, but starvation alone is insufficient to explain the lack of recovery and increased mortality with time (14). Continuous contamination of the aquatic environment with neonicotinoids may undermine the invertebrate resource base of aquatic ecosystems (11), thereby indirectly reducing populations of fish, birds, bats, frogs, and other animals that feed on them. Indeed, the steady decline of five species of birds in the Netherlands over the past two decades correlates with imidacloprid contamination of surface waters during the same period (2).

The effects of neonicotinoid residues on soil biota remain largely unknown, but the extreme efficiency with which these insecticides eliminate grub populations in turf is worrisome (4). This issue requires more study, because the ecosystem services provided by soil organisms are essential for sustainable agricultural production (15). Scattered seeds coated with high concentrations of neonicotinoids may also pose a risk to birds and rodents, despite the higher tolerance of vertebrates due to their distinct nicotinic receptor subunits (4).

Mechanisms that underpin chronic neonicotinoid effects on terrestrial and aquatic arthropods include immune suppression and feeding inhibition. While these and other issues are investigated further, current knowledge calls for a reconsideration of current prophylactic seed treatments with neonicotinoids. Such treatments are the main source of soil and water contamination; are often unnecessary, as they either do not increase yields or are not profitable; and go against the principles of integrated pest management (15). ■



Fate of neonicotinoids and pathways of environmental contamination.

lack of effect. The amount of honey produced is usually higher in contaminated hives because feeding inhibition and death of workers result in excess honey stores. Also, some undetectable sublethal effects cause mortality after a time lag (1). Finally, honey bee colonies compensate forager losses by producing hundreds of new workers daily; colonies thus usually overcome the initial effects during spring and summer and may survive the winter apparently unscathed. However, colony growth is usually hampered by queen failure in the next season (7), indicating that the queen suffers the effects of long-term intoxication. Bumble bees produce 85% fewer queens per colony when exposed to field-realistic concentrations of imidacloprid (8).

Concurrent with the widespread use of neonicotinoids, honey bees have experienced an increase in viral diseases, some of which are propagated by a mite parasite (*Varroa destructor*) that undermines bee health. Pollen from monoculture crops also

insect larvae, particularly mayflies, caddisflies, stoneflies, and midges. They are also toxic to decomposer amphipods, woodlice, and most crustaceans, but water fleas are very tolerant (4).

Because most neonicotinoids persist in soils for a year or more and are water soluble, 80 to 98% of residues remaining in the soil of treated crops eventually move into surface waters or leach into groundwater. Recent surveys from nine countries show 80% of surface waters contaminated with neonicotinoids at levels of 0.14 to 18 ppb, which are sublethal to aquatic arthropods (10, 11). However, as in bees, chronic toxicity in all these organisms involves delayed and cumulative lethal effects over time (12). Experiments in aquatic model ecosystems treated with single or repeated dosages of imidacloprid confirm this: midges, ostracods, and mayflies disappear; their populations do not recover while residues in water are above 1 ppb (13). After 8 years of field monitoring, Van Dijk *et al.* (11) reported that imidacloprid concentrations as low as 0.01 ppb led to significant reduction of macroinvertebrates in surface waters.

## REFERENCES

1. G. Di Prisco *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18466 (2013).
2. C. A. Hallmann *et al.*, *Nature* **511**, 341 (2014).
3. F. Sánchez-Bayo, K. Goka, *PLOS ONE* **9**, e94482 (2014).
4. L. W. Pisa *et al.*, *Environ. Sci. Pollut. Res.*, 10.1007/s11356 (2014).
5. A. Decourtye, J. Devillers, in *Advances in Experimental Medicine and Biology: Insect Nicotinic Acetylcholine Receptors*, S. H. Thany, Ed. (Landes Bioscience, Austin, TX, 2009), pp. 85–95.
6. F.-X. Dechaume Moncharmont, A. Decourtye, C. Hennequet-Hantier, O. Pons, M. H. Pham-Delégue, *Environ. Toxicol. Chem.* **22**, 3088 (2003).
7. C. Sandrock *et al.*, *PLOS ONE* **9**, e103592 (2014).
8. P. R. Whitehorn, S. O'Connor, F. L. Wackers, D. Goulson, *Science* **336**, 351 (2012).
9. G. Rondeau *et al.*, *Sci. Rep.* **4**, 5566 (2014).
10. A. R. Main *et al.*, *PLOS ONE* **9**, e92821 (2014).
11. T. C. Van Dijk *et al.*, *PLOS ONE* **8**, e62374 (2013).
12. M. A. Beketov, M. Liess, *Environ. Toxicol. Chem.* **27**, 461 (2008).
13. D. Hayasaka *et al.*, *Ecotoxicol. Environ. Saf.* **80**, 355 (2012).
14. A.-M. Nyman *et al.*, *PLOS ONE* **8**, e62472 (2013).
15. M. Chagnon *et al.*, *Environ. Sci. Pollut. Res.*, 10.1007/s11356 (2014).

Faculty of Agriculture and Environment, The University of Sydney, Eveleigh, NSW 2015, Australia. E-mail: sanchezbayo@mac.com



## HUMAN GENETICS

# Hidden features of human hotspots

Multiple flavors of genetic changes occur at meiotic recombination sites

By Bernard de Massy

During the production of gametes (meiosis), genetic information is exchanged between chromosomes by a process called homologous recombination. Recombination is initiated by the formation of DNA double-strand breaks (DSBs) (1). The repair of DSBs generates reciprocal (or crossing over) and nonreciprocal (non-crossing over) exchanges between chromosomes (2). In humans and mice, recombination occurs at specific sites along chromosomes, also called hotspots, which are predicted binding sites for PR domain containing 9 (PRDM9) protein (3, 4). Sequence variations (polymorphisms) in PRDM9 can lead to differences in crossing-over localization. On page 826 of this issue, Pratto *et al.* (5) reveal that recombination activity along chromosomes varies among individuals and uncover several unexplored properties of meiotic recombination in humans.

One major challenge in studying meiotic recombination is to characterize not only crossing-over events, but ideally all recombination activity. Indeed, crossing over represents only a minor fraction (about 10%) of meiotic recombination events, and most DSBs are thought to give rise to gene conversion without crossing over. To address this issue, an approach developed in mice (6, 7) detects DSB repair intermediates in which the recombinase DMCI1 is bound to single-strand DNA. The data obtained by high-throughput sequencing of the purified DNA fragments are then processed by extensive bioinformatics analyses. An intrinsic advantage of this strategy is the possibility to detect meiotic recombination activity in the absence of genetic markers. A limitation is that intermediates are by definition transient and thus their detection depends on their life span.

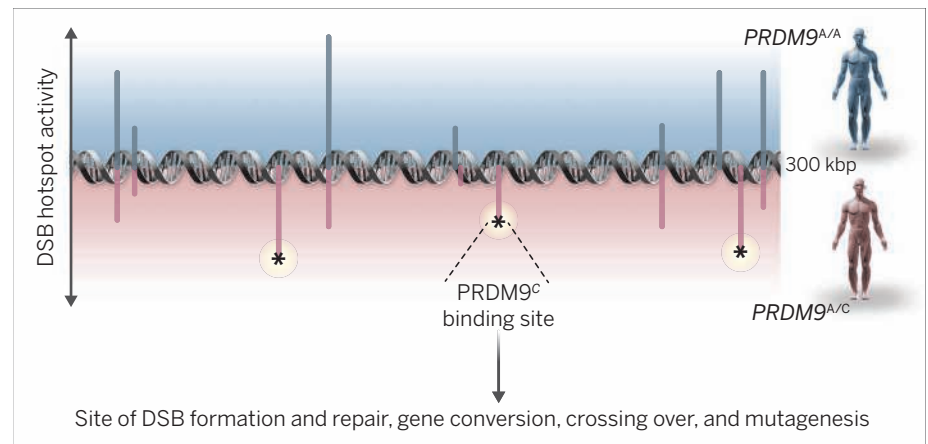
With this method, Pratto *et al.* mapped DSB repair intermediates in an individual homozygous for the major PRDM9 allele (PRDM9<sup>A/A</sup>) and in a heterozygous individual (PRDM9<sup>A/C</sup>). They show that about half of the hotspots (around 19,000 sites) of the PRDM9<sup>A/C</sup> individual do not overlap with those of the PRDM9<sup>A/A</sup> individual, revealing the genome-wide control of hotspot

localization mediated by PRDM9 (see the figure). This extends the previous finding that variation of crossing-over and non-crossing-over activity at individual hotspots correlates with polymorphisms of PRDM9 in humans (8).

An intriguing feature of human meiotic recombination, which had been partially described by high-resolution crossover localization derived from linkage disequilibrium maps (the map of association between genes along chromosomes based on analysis of genetic variation in a population), is the specific patterns of nucleotide substitutions and polymorphisms around hotspots (9). The high-resolution map of

distributed around hotspot centers with higher G-to-A and T-to-C transitions on one side and C-to-T and A-to-G on the other side [adenine (A), guanine (G), thymine (T), cytosine (C)]. DNA synthesis taking place on both sides of the DSB upon repair may be responsible for this effect, which could have a noticeable evolutionary impact on genome diversity.

Another feature of human hotspots is a local enrichment for substitutions toward GC, the extent of which is correlated with recombination activity. This could be explained by a process whereby gene conversion events are biased toward GC alleles (10). Evidence for such bias has been



**Variations and hotspots.** Meiotic recombination sites (hotspots) in a 300-kbp region of the genome of two individuals with distinct PRDM9 genotypes (A/A and A/C) only partially overlap. Hotspots specific to the individual with the A/C genotype have a DNA motif predicted to be recognized by the PRDM9<sup>C</sup> variant (labeled with an asterisk). Meiotic DSB repair can lead to gene conversion with (crossing over) or without reciprocal exchange (non-crossing over). DSB repair is also associated with increased mutation rate and biased gene conversion toward GC within 1 to 2 kbp around the hotspot center.

recombination intermediates reveals additional aspects of localized nucleotide diversity within 1 to 2 kilobase pairs (kbp) around hotspots. For example, an asymmetric purine-pyrimidine skew detected around hotspot centers appears to be an intrinsic hotspot feature and not a consequence of recombination activity. Indeed, it is also detected in orthologous sequences of the chimpanzee genome, which uses different sites of recombination due to the distinct DNA binding specificities of chimpanzee PRDM9 alleles.

Analysis of recent polymorphisms in the human lineage also highlights an unexpected pattern, best explained by an increased mutation rate, asymmetrically

demonstrated at two human hotspots (11). The underlying molecular mechanism is not known and may be due to biased repair of heteroduplex DNA during DSB repair. Such bias has an appreciable impact on GC content evolution, but its biological implications remain to be determined. Understanding the long-term consequences of these features will require integrating the dynamic properties of recombination hotspots caused by the erosion of PRDM9 binding sites due to gene conversion (4).

At low frequency, DSB repair also could occur by homologous recombination between nonallelic sequences [nonallelic homologous recombination (NAHR)], typically a sequence homologous to the region

of DSB and that is repeated elsewhere in the genome. Such DSB repair can lead to genome rearrangements that are responsible for several human pathologies (12). The high-resolution maps of recombination intermediates of Pratto *et al.* revealed that 14 out of 27 rearrangements due to NAHR—the breakpoints of which were mapped within less than 10 kbp—overlap with a meiotic recombination site, suggesting that these events have a meiotic origin and depend on *PRDM9*. This finding extends the discovery that the rearrangement leading to Charcot-Marie-Tooth type 1A disease, a hereditary motor and sensory neuropathy, depends on *PRDM9* (8). Thus, the occurrence of rearrangements due to NAHR is predicted to be influenced by the *PRDM9* genotype, an observation that will be important for future research on these human diseases.

The approaches developed for analyzing meiotic recombination and the data obtained by Pratto *et al.* provide exciting perspectives in terms of molecular mechanisms and genome evolution. Particularly, non-crossover events, an important outcome of meiotic DSB repair, need further exploration. Analysis of recombination activity in repeated regions of the genome that are inherently difficult to investigate will be a major challenge. The study by Pratto *et al.* also establishes several important parameters for hotspot evolution that need to be quantified to understand hotspot dynamics and *PRDM9* activity and the consequences of meiotic recombination activity on genome diversity. These approaches could potentially be applied also to investigate meiotic recombination activities in other species. ■

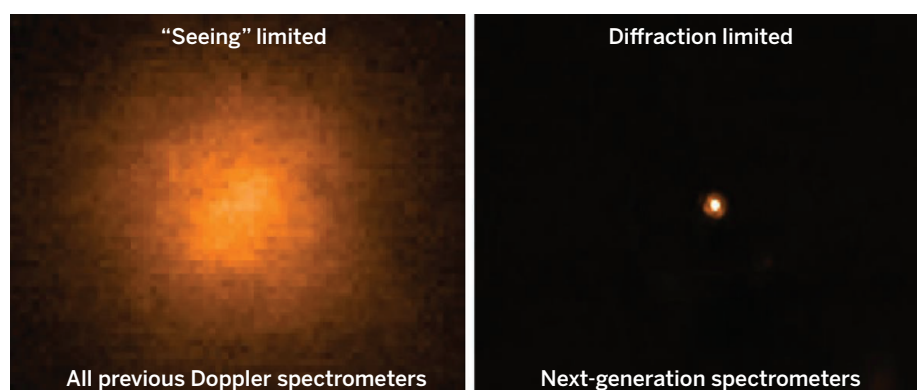
#### REFERENCES AND NOTES

1. B. de Massy, *Annu. Rev. Genet.* **47**, 563 (2013).
2. F. Baudat, Y. Imai, B. de Massy, *Nat. Rev. Genet.* **14**, 794 (2013).
3. F. Baudat *et al.*, *Science* **327**, 836 (2010).
4. S. Myers *et al.*, *Science* **327**, 876 (2010).
5. F. Pratto *et al.*, *Science* **346**, 1256442 (2014).
6. F. Smagulova *et al.*, *Nature* **472**, 375 (2011).
7. K. Brick, F. Smagulova, P. Khil, R. D. Camerini-Otero, G. V. Petukhova, *Nature* **485**, 642 (2012).
8. I. L. Berg *et al.*, *Nat. Genet.* **42**, 859 (2010).
9. A. Auton *et al.*, *Science* **336**, 193 (2012).
10. L. Duret, N. Galtier, *Annu. Rev. Genomics Hum. Genet.* **10**, 285 (2009).
11. L. Odenthal-Hesse, I. L. Berg, A. Veselis, A. J. Jeffreys, C. A. May, *PLOS Genet.* **10**, e1004106 (2014).
12. P. Liu, C. M. Carvalho, P. J. Hastings, J. R. Lupski, *Curr. Opin. Genet. Dev.* **22**, 211 (2012).

#### ACKNOWLEDGMENTS

B.d.M. is supported by the Centre National de la Recherche Scientifique and the European Research Council Executive Agency under the European Community's Seventh Framework Programme (FP7/2007-2013 grant agreement no. 322788).

10.1126/science.aaa0612



**Seeing-limited images versus diffraction-limited images.** An AO-fed spectrometer achieves much higher spatial resolution, and hence spectral resolution, compared to presently available Doppler instruments, owing to fundamental changes in its design. All such design modifications lead to improved velocity precision. Figure adapted from (12).

#### ASTRONOMY

## Improving planet-finding spectrometers

Adaptive optics will enhance the precision of exoplanet surveys

By Justin R. Crepp

**A**daptive optics (AO) systems correct for optical wavefront errors introduced by Earth's turbulent atmosphere, turning initially blurry images into intense diffraction-limited concentrations of light. The implementation of AO systems on the world's largest telescopes has revolutionized essentially all areas of astronomy (1). Instruments that receive a well-corrected beam of light can operate as if observing from space and thus benefit from an order-of-magnitude higher spatial and spectral resolution (see the figure). Given the benefits of working with nonfuzzy images, it may therefore be surprising to learn that one of the most important techniques for finding extrasolar planets, the Doppler radial velocity method, still uses "seeing-limited" observations—that is, measurements obtained without AO correction.

Doppler instruments began detecting extrasolar planets in the mid-1990s, well before AO systems became commonplace at astronomical facilities. Thus, radial velocity measurements were by default seeing limited (2). The targeting of Sun-like stars in search of solar-system analogs and development of the iodine gas cell for velocity calibration (which only operates in the 0.50- to 0.62- $\mu\text{m}$  wavelength range) further reinforced any initial "Copernican-esque" tendencies to observe at visible wavelengths, thus precluding diffraction-limited capabilities from the ground.

The Doppler method has since matured, and observations have naturally branched out to study planets orbiting different types of stars (3). For example, we now know that the lowest-mass stars in the galaxy, the M-dwarfs, have a propensity to form terrestrial planets (4). NASA's Kepler spacecraft has corroborated this result, finding that many of these worlds orbit near the habitable zone (5). As such, several teams are now developing the first precision infrared spectrometers to gain access to the black-body radiation of M-dwarfs, which peaks at wavelengths in the 1- to 2- $\mu\text{m}$  range (6, 7).

Although the scientific motivation for building radial velocity instruments for the infrared is clear, an important threshold crossing occurs when observing at wavelengths just beyond 1.0  $\mu\text{m}$ : the ability to achieve diffraction-limited performance by correcting for atmospheric turbulence using AO. Instruments designed to operate behind an AO system have fundamentally different properties than their seeing-limited counterparts. In particular, their physical size shrinks commensurately with the size of the input image. It turns out that all such hardware modifications resulting from this effect help to improve velocity precision.

For example, one of the most problematic sources of error encountered by present-

Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA. E-mail: jcrepp@nd.edu



day Doppler instruments involves their stability. Precise velocity time-series measurements require a spectrometer with exquisite temperature and pressure control. Any changes within the instrument's optical path (tiny movements of lenses, mirrors, gratings, filters, etc.) can mimic a Doppler shift. To detect the signal of an Earth-like planet orbiting in the habitable zone, thermal fluctuations must be reduced to the millikelvin level (8).

The size of a seeing-limited spectrometer increases with telescope diameter. For the 8- to 10-m-diameter telescope facilities, it is not uncommon for such instruments to be the size of a small car, making precise environmental control challenging. However, a variety of creative control solutions exist to overcome this challenge for spectrometers operating at the diffraction limit. For instance, instruments that are an order of magnitude smaller in all directions can be given appreciable thermal inertia by enclos-

### ***"The singular challenge to realizing a diffraction-limited Doppler instrument involves coupling starlight into small (single mode) optical fibers."***

ing them in a bath of liquid cryogenics. Also, more exotic materials may be used, such as invar—an iron-nickel alloy having a coefficient of thermal expansion a factor of 10 lower than that of the aluminum-based materials normally used for astronomical instruments. Further, a compact optical design is less expensive, benefits from components with higher optical quality, has a faster development time scale, and can achieve an ultrahigh vacuum.

Diffraction-limited spectrometers also solve a host of other problems in addition to stability. Background light contamination from the Moon, the sky, and unseen neighboring stars is reduced by three orders of magnitude. Modal noise from fiber optic cables, introduced by changes in the interference pattern of starlight, can be eliminated entirely by using single-mode fibers. And, by working in the infrared, astrophysical jitter caused by star spots that rotate in and out of view as a star spins is reduced compared to that in visible-light instruments. These systematic effects dominate the error budgets of present-day radial velocity spectrometers.

The singular challenge to realizing a diffraction-limited Doppler instrument

involves coupling starlight into small (single mode) optical fibers. This barrier has recently been overcome by Jovanovic *et al.* (9). Using a new "extreme" AO system built for the Subaru 8.2-m telescope, single-mode fiber coupling efficiencies of 68% have been demonstrated with an 8- $\mu$ m-diameter fiber core, approaching the theoretical maximum attainable limit set by the level of delivered wavefront correction. The observations were acquired in broadband light for wavelengths near 1.55  $\mu$ m, just longwards of the ideal 1.0- to 1.1- $\mu$ m window (the Y-band) where the Earth's atmosphere is uniformly transmissive and not riddled with OH-emission lines (10). The implications of these measurements are that Doppler spectrometers the size of a shoebox are feasible and that instrument throughput will be sufficient to permit record-breaking precision with existing telescope facilities.

Indeed, a substantive improvement in velocity precision would be timely. NASA's forthcoming Transiting Exoplanet Survey Satellite will survey the entire night sky, revealing the addresses of several hundred nearby rocky planets that may resemble Earth (11). Most of the terrestrial worlds found to reside near the habitable zone will induce Doppler reflex motions well below the detection limits of presently available spectrometers.

Next-generation Doppler instruments promise to be smaller and more powerful than their seeing-limited counterparts. Despite being optimized for near-infrared wavelengths, diffraction-limited spectrometers will not only be ideally suited for M-dwarf stars, but can also study nearby Sun-like stars with unprecedented detail. The errors that conspire to overwhelm the Doppler signal of an Earth analog can be mitigated by developing instruments that benefit from the unimpaired vision of telescopes that correct for atmospheric turbulence using adaptive optics. ■

#### REFERENCES

1. R. Davies, M. Kasper, *Annu. Rev. Astron. Astrophys.* **50**, 305 (2012).
2. M. Mayor, D. Queloz, *Nature* **378**, 355 (1995).
3. J. Johnson, K. M. Aller, A. W. Howard, J. R. Crepp, *Publ. Astron. Soc. Pac.* **122**, 905 (2010).
4. X. Bonfils *et al.*, *Astron. Astrophys. A* **549**, A109 (2013).
5. J. Johnson, *Phys. Today* **67**, 31 (2014).
6. A. Quirrenbach *et al.*, *Proc. SPIE* **8446**, 84460R (2012).
7. S. Mahadevan *et al.*, *Proc. SPIE* **8446**, 84461S (2012).
8. C. Lovis, D. Fischer, in *Exoplanets*, S. Seager, Ed. (Univ. of Arizona Press, Tucson, AZ, 2011), p. 526.
9. N. Jovanovic *et al.*, *Proc. SPIE* **9147**, 287 (2014).
10. S. Redman, thesis, The Pennsylvania State University (2011).
11. G. Ricker *et al.*, <http://arxiv.org/abs/1406.0151> (2014).
12. H. Takami, *Proc. SPIE* **7014**, 701404 (2008).

10.1126/science.1262071

#### DEVELOPMENTAL BIOLOGY

## ***Mesenchymal progenitor panoply***

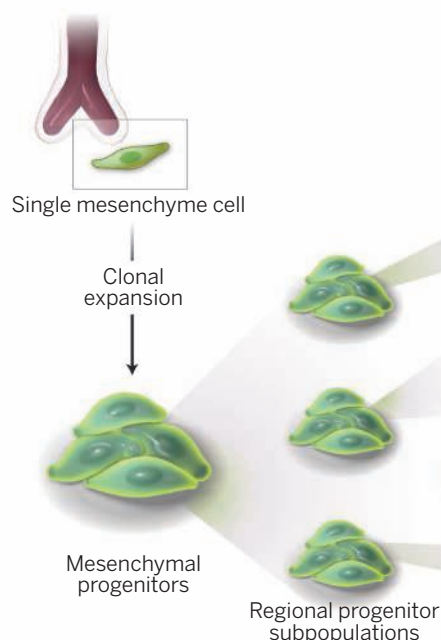
Distinct pools of progenitor cells develop stromal tissues of the lung

By Joo-Hyeon Lee and Carla F. Kim

Lung development relies on temporal and spatial cues to pattern the structures that sustain life through oxygen exchange. Signals from mesenchymal cells that develop into the stromal tissues of the lung are key to this elaborate process (1, 2). How the epithelial cells lining the respiratory tract arise from progenitor cells is better understood than the differentiation of mesenchymal progenitor cells. On page 827 of this issue, Kumar *et al.* (3) define dynamic niches that regulate multiple mesenchymal progenitor cell pools during mammalian lung development. The study suggests that mesenchymal progenitor cells are diverse and differentiate in response to distinct microenvironmental cues.

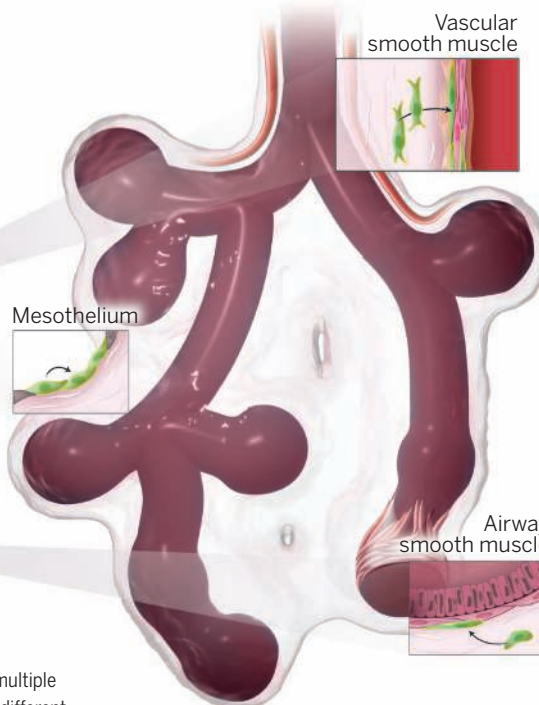
The primitive airway epithelium of the murine lung is specified at embryonic day 9.5 (E9.5) and undergoes a process of bud formation and branching morphogenesis while surrounded by mesenchymal progenitors (4). During this process, the distal buds of the airway contain multipotent epithelial progenitor cells that expand and leave behind lineage-restricted airway epithelial progenitors (5). Kumar *et al.* show that neighboring mesenchymal cells differentiate as development proceeds from E10.5 to E14.5. The authors used two different strategies in the mouse involving clonal cell labeling and the expression of colored reporter proteins for lineage tracing at the single-cell level, akin to clonal analysis traditionally done in *Drosophila*. Although Kumar *et al.* show that mesenchymal progenitor cells migrate and intermix, left-right boundaries of the lung were already established prior to cell labeling at E9. This suggests that mesenchymal cells segregate to left versus right lobes of the lung prior to budding. The finding also raises the possibility that mesenchymal progenitors have a more active role in early lung morphogenesis than previously appreciated.

## Early lung bud



**Niche influences.** During early mouse lung development, multiple subpopulations of mesenchymal progenitor cells form. The different microenvironments influence their differentiation into specific mesenchymal cells.

## Developing lung



The clonal analyses by Kumar *et al.* complement our current understanding of how mesenchymal cells differentiate and organize during lung development. This patterning process is of considerable interest as mesenchymal cells are proving to be important in maintaining tissue homeostasis and in disease. The authors observed that early in development, mesenchymal progenitor cells destined to establish the lung's mesothelium (a thin protective lining covering the lung) were separate from the other mesenchymal progenitor pools. Later in development, pools of cells emerged that contained mixes of progenitors destined to become fibroblasts, mesothelial cells, and other mesenchymal cell types. At E13.5, the majority of the cell pools (clones) contained only one mesenchymal progenitor cell type. These were in discrete locations (see the figure); for example, clones containing progenitors for vascular smooth muscle cells were observed near developing vasculature of the lung. Indeed, a previous study showed that cells in the developing mouse lung that express fibroblast growth factor 10 (which is required for lung morphogenesis) retain mesenchymal progenitor pools for multiple mesenchymal lineages (6). Multiple mesenchymal cell types, including smooth muscle cells, have also been shown to be derived from a multipotent cardiopulmonary progenitor (7).

Kumar *et al.* emphasize that mesenchymal progenitor cells appear to be maintained in an undifferentiated state until proximity defines the lineage potential in the mesenchymal compartment. Signals released by the vasculature or airway influence the type of differentiation that will occur. The findings are consistent with the "sandwich model" for patterning lung mesenchyme (8). In this model, mesenchyme located between mesothelium and airways is in an undifferentiated state until the closest signals promote differentiation.

Focusing on airway smooth muscle, Kumar *et al.* defined the mesenchymal progenitor niche neighboring the distal tips of the airways. Their observations suggest that as lung development proceeds, mesenchymal progenitor cells divide and the daughter cells differentiate as they move upwards along the proximal-distal axis. These cells lose progenitor capacity and become lineage restricted to the airway smooth muscle fate. They cannot revert back into a progenitor lineage with the formation of a new branch point. Instead, new undifferentiated stalk mesenchymal cells are recruited to newly forming branch domains, implying that there is a lineage boundary between airway smooth muscle cells and airway lateral stalk cells.

The process of recruitment and differentiation into the airway smooth muscle lineages is governed in part by local signaling by the secreted molecule Wnt. Genetic mouse models demonstrated the importance of global

Wnt signaling in specifying epithelial and mesenchymal progenitors (9). Kumar *et al.* expand these findings by demonstrating the role of Wnt signaling in enabling stalk cells to become tip progenitors that can move up the airway. Thus, the establishment of tip mesenchymal progenitors requires specific signaling from neighboring cells. Indeed, altered Wnt signaling impaired vascular and airway smooth muscle patterning in mice. The findings suggest how epithelial and endothelial cells are not sufficient for directing the differentiation of mesenchymal progenitor pools, yet may still play a role in fine tuning their fate.

The study of Kumar *et al.* raises new questions about mesenchymal identities and the cell-cell interactions between neighboring epithelial, endothelial, and mesenchymal cells in the lung. It is still not clear how the locations of mesenchymal progenitor pools are determined. Whether

a lineage boundary is defined by active signals such as growth factors, or by a passive process based on principles of gradient and diffusion, also remains to be determined. It may be that mesenchymal progenitors are reactivated in response to injury or in lung disease. Particularly relevant is the question of whether mesenchymal progenitors are altered in asthma, a condition that affects smooth muscle functions. If so, understanding mesenchymal differentiation and patterning may reveal new therapeutic avenues for asthma and other lung conditions. The necessity and sufficiency of mesenchymal progenitor populations in epithelial development also remains to be determined. Genomic and transcription analyses at the single-cell level should increase our understanding of this panoply of lung mesenchymal lineages and may help discern whether other organs use similar mesenchymal progenitor pools. ■

## REFERENCES AND NOTES

1. E. E. Morrissey, B. L. Hogan, *Dev. Cell* **18**, 8 (2010).
2. D. Warburton *et al.*, *Curr. Top. Dev. Biol.* **90**, 73 (2010).
3. M. E. Kumar *et al.*, *Science* **346**, 1258810 (2014).
4. W. V. Cardoso, J. Lü, *Development* **133**, 1611 (2006).
5. E. L. Rawlins, C. P. Clark, Y. Xue, B. L. Hogan, *Development* **136**, 3741 (2009).
6. E. El Agha *et al.*, *PLOS ONE* **7**, e38452 (2012).
7. T. Peng *et al.*, *Nature* **500**, 589 (2013).
8. M. Weaver, L. Batts, B. L. Hogan, *Dev. Biol.* **258**, 169 (2003).
9. E. D. Cohen *et al.*, *J. Clin. Invest.* **119**, 2538 (2009).

## ACKNOWLEDGMENTS

We thank X. Sun and J. Mahoney for helpful discussions.

10.1126/science.aaa0609

Children's Hospital, Harvard Medical School, Boston, MA, USA.  
E-mail: carla.kim@childrens.harvard.edu



## PALEOCLIMATE

# Did Antarctica initiate the ice age cycles?

Gradual buildup of Antarctic ice sheets preceded extensive Northern Hemisphere glaciation about 3 million years ago

By Robert McKay

**A**t the end of the Pliocene epoch, about 2.6 million years ago, a sustained period of global cooling coincided with the growth of continental ice sheets over Europe and North America. Since then, these ice sheets have periodically grown and collapsed more than 50 times, driving global sea level fluctuations of more than 100 m. What caused this bipolar glacial world? On page 847 of this issue, Woodard *et al.* (1) show that the ice sheets began to grow gradually before 2.7 million years ago. This gradual buildup led to a more efficient heat transport between the Atlantic and Pacific abyssal oceans, thus resulting in an Atlantic cooling.

The authors first measured magnesium/calcium and oxygen isotope ratios in the calcium carbonate shells of bottom-dwelling organisms called foraminifera. They then used these ratios to reconstruct changes in ocean temperature and global ice volume. The results suggest a progressive increase in global ice volume, with the equivalent of ~20 m of sea level locked up in permanent ice sheets in Earth's polar regions between 3.1 and 2.7 million years ago (1). This increase in ice volume preceded the first compelling evidence for major Northern Hemisphere glaciation across North America around 2.7 million years ago (2); the authors thus infer that

most of this ice buildup occurred in the Antarctic. There is some evidence for glaciation in Greenland and the Arctic margin between 3.1 and 2.7 million years ago; it is unlikely, however, that this glaciation dominated this permanent-ice signal, considering that only ~7 m of sea level equivalent ice currently lies on Greenland [versus ~58 m in Antarctica, ~22 m of which is grounded below sea level (3)].

Other studies provide support for the conclusion that Antarctic ice volumes were comparatively low in the Early Pliocene and that their growth in the Late Pliocene could have been the main contributor to the permanent-ice buildup at this time. A core collected by the ANDRILL Program has revealed cycles of growth and retreat of the West Antarctic Ice Sheet during the Pliocene (4). However, this ice sheet probably contributed no more than 3 m to the permanent-ice growth between 3.1 and 2.7 million years ago (5). Expedition 318 of the Integrated Ocean Drilling Program provided compelling evidence that the ice may instead have been located in the marine-based sectors of the giant East Antarctic Ice Sheet. The geochemistry of marine muds offshore of East Antarctica indicates that rocks currently hidden under the ice sheet in the Wilkes subglacial basin were eroded by an ice sheet that had retreated substantially inland in the early Pliocene (6). This marine-based East Antarctica retreat potentially contributed to sea level rises of ~20 m above present-day levels during the Early Pliocene interglacials (7). Late Pliocene cooling and

expansion of the West Antarctic Ice Sheet is documented in the ANDRILL core between 3.3 and 2.6 million years, and it is likely that the marine-based sectors of the East Antarctic Ice Sheet also expanded at this time (4, 8).

Woodard *et al.* show that the Late Pliocene rise in global ice volume was followed by a rapid ~2°C cooling of North Atlantic deep-water temperatures and a ~1.5°C increase in North Pacific deep-water temperatures. Thus, Antarctic ice sheet expansion coincided with the disappearance of a ~4°C temperature gradient that previously existed between the bottom waters of the two ocean basins, indicating more efficient heat transport between the deep North Atlantic and the Pacific Ocean.

The authors argue that a major Antarctic ice sheet advance was likely associated with sea ice expansion and enhanced ocean stratification, which would have restricted upwelling of warm, salty North Atlantic-sourced water in the Southern Ocean. Heat transported south by the Atlantic's meridional overturning circulation could thus not be as readily released to the atmosphere around Antarctica and was instead carried into the abyssal Pacific Ocean. These changes in heat distribution may have been substantial enough on their own to initiate Northern Hemisphere glaciation, but other important feedbacks related to Antarctic cooling may also have come into play.

A decline in opal productivity in the high-latitude Southern Ocean is attributed to declining light levels or nutrient availability

Antarctic Research Center, Victoria University of Wellington, Wellington, New Zealand. E-mail: robert.mckay@vuw.ac.nz

## Pliocene Antarctica



## Present-day Antarctica



**Antarctica's marine-based ice buildup.** Woodard *et al.* suggest that permanent ice cover built up in Antarctica between 3.1 and 2.7 million years ago. This ice cover would have fundamentally altered key Southern Ocean feedbacks that may have contributed to the glacial onset in the Northern Hemisphere.

supplied by upwelling water masses at 2.7 million years ago. Either of these scenarios points to enhanced stratification and sea ice coverage after this time (9, 10). Glacial sediments, diatom-based proxies, and lipid biomarker-based proxies in the ANDRILL record also indicate a prolonged cooling and changing stratification of the surface waters at the Antarctic margin beginning 3.3 million years ago (8). This capping of the Southern Ocean may have led to numerous other possible feedbacks that could have contributed to Late Pliocene cooling. For example, reduced ventilation of the deep waters around Antarctica may have caused more atmospheric CO<sub>2</sub> to be sequestered in the abyssal ocean. Indeed, paleorecords indicate that atmospheric CO<sub>2</sub> levels fell from above 400 ppm to below 280 ppm in the Late Pliocene, with the sharpest decline near 2.7 million years ago (11)—a threshold considered key for the triggering of the Northern Hemisphere glaciation (12). Also, a steepened pole-to-equator temperature gradient resulting from the Antarctic cooling may have led to a contraction of the Southern Hemisphere subtropical gyres, reducing heat and salt exchange between the major ocean basins by surface ocean currents (8).

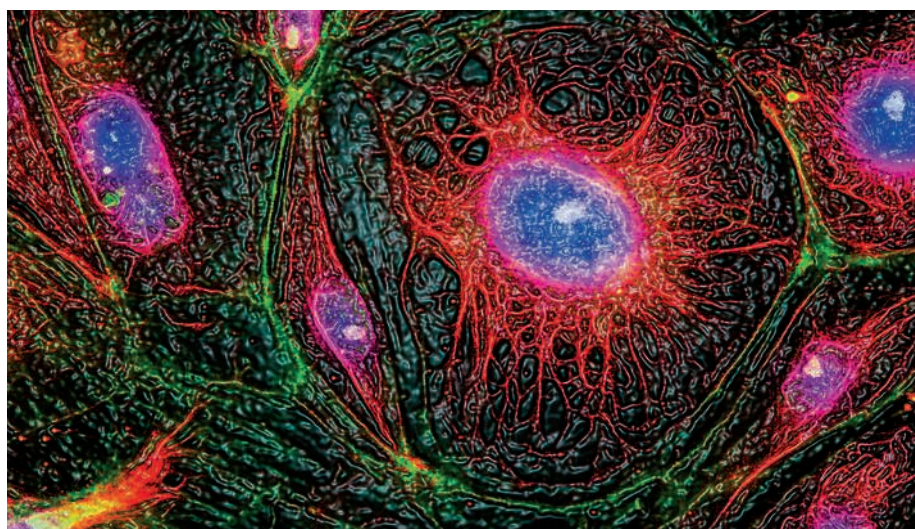
Other suggested mechanisms for the Late Pliocene cooling involve wholesale shifts in Northern Hemisphere oceanic and atmospheric circulation patterns driven by tectonic events, or low-latitude processes such as the termination of hypothesized permanent El Niño conditions in the Pacific. Although tectonic events may indeed be the ultimate driver for the onset of Northern Hemisphere ice ages, it is becoming increasingly clear that a major cooling step occurred in the Antarctic before 2.7 million years ago. Given the importance of the Southern Ocean for the carbon cycle and the redistribution of heat around the planet, mechanisms to help explain the cooling of the Antarctic in the broader context of Late Pliocene climate should be explored further. ■

#### REFERENCES AND NOTES

1. S. C. Woodard et al., *Science* **346**, 847 (2014).
2. I. Bailey et al., *Quat. Sci. Rev.* **75**, 181 (2013).
3. P. Fretwell et al., *Cyrosphere* **7**, 375 (2013).
4. T. Naish et al., *Nature* **458**, 322 (2009).
5. D. Pollard, R. M. DeConto, *Nature* **458**, 329 (2009).
6. C. P. Cook et al., *Nat. Geosci.* **6**, 765 (2013).
7. K. G. Miller et al., *Geology* **40**, 407 (2012).
8. R. McKay et al., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6423 (2012).
9. D. M. Sigman et al., *Nature* **428**, 59 (2004).
10. C.-D. Hillenbrand, G. Cortese, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **242**, 240 (2006).
11. IPCC, *Climate Change 2013: The Physical Science Basis* (Cambridge Univ. Press, Cambridge, 2013).
12. R. M. DeConto et al., *Nature* **455**, 652 (2008).

#### ACKNOWLEDGMENTS

I thank T. Naish for discussion and the Rutherford Discovery Fellowship (RDF-13-VUW-003) for support.



Synthetic circuits program living cells to remember their past.

#### SYNTHETIC BIOLOGY

## Dynamic genome engineering in living cells

Engineered gene expression systems provide cells with a molecular memory of their past

By Simon Ausländer and Martin Fussenegger

Living cells continuously measure, process, and store cellular and environmental information in response to specific signals. Bioengineers are now starting to use these systems to build customized genetic regulatory circuits that can control targeted biological processes. They are also using them to develop new regulatory modes and novel biochemical pathways (1). In combination with new genome-editing tools (2), this technology holds great promise for the development of biomedical and biotechnological applications of specially engineered “designer” cells. On page 825, Farzadfar and Lu (3) move us closer to this goal by constructing a cellular memory device that is based on a conditional gene-editing platform. In this way they have gained access to the enormous storage capacity of genomic DNA to record analog information.

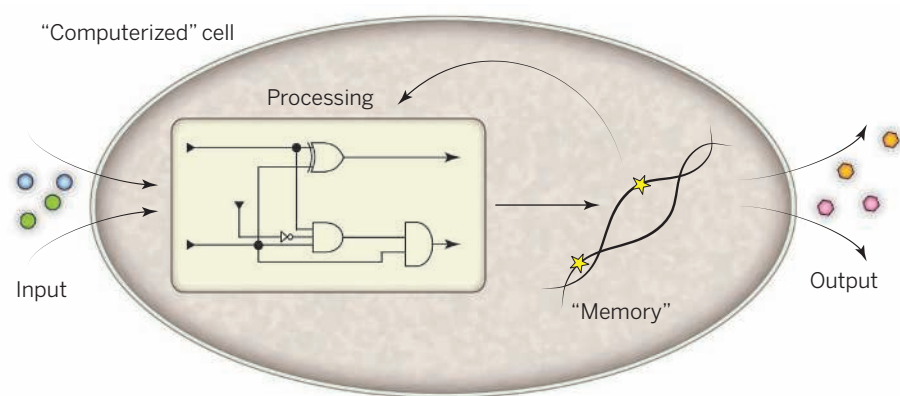
The design of reliable and robust gene circuits faces a number of challenges, including the need for (i) a library of well-characterized gene switches that can be scaled up into complex regulatory circuits; (ii) the stable, long-term propagation of multicircuit networks that ensure correct and dynamic functionality in the cell; and (iii) synthetic

memory devices that guarantee sustained storage of the information that is computed by the genetic program. The first synthetic memory devices were based on two distinct small molecule-responsive gene expression units that were configured for mutually exclusive transcription control. These negative feedback loops enable reversible switching of bistable reporter protein expression by a specific trigger compound (4, 5).

More recently, bacteriophage-derived recombinases, which can bind and cut DNA sequences flanked with specific recognition sites, have advanced the design of synthetic on-off circuits (6, 7). Different recombinase-specific sites can be used in combination to control reporter gene expression, thus simplifying circuit design. Moreover, these recombinase-based genetic circuits can perform complex logic functions (AND, OR) and have been shown to record and reset corresponding target gene expression states (6). The number of available recombinases, however, limits overall circuit complexity, and the approach is laborious, requiring the integration of each recombinase recognition site at the appropriate DNA target locus.

Farzadfar and Lu neatly sidestep this problem and at the same time harness the entire genomic DNA to store synthetic circuit-computed biological information. Instead of using conventional recombinases





**Controlling “computerized” cells.** Synthetic gene circuits perform information processing tasks that include three main processes. Environmental and cellular signals (inputs) are detected by biosensors (e.g., trigger-responsive transcription factors) that transmit the information into the gene circuit for interpretation and logic processing. The circuit’s architecture determines the executed computation and controls the output response. When coupled to a synthetic memory device, the computed output information can be stored in genomic DNA that in turn can feed again into the circuit to modify the circuit’s behavior.

that rely on site-specific recognition, the authors took advantage of the beta recombinase from bacteriophage  $\lambda$ , which uses a single-stranded DNA (ssDNA) oligonucleotide to target complementary DNA and mediate recombination, thereby enabling precise genome editing. While the oligonucleotide template encoding the desired DNA modification is normally exogenously delivered to the cells, the method of Farzadfard and Lu enables the efficient production of ssDNA templates in living bacteria. The approach is based on a bacterial-derived “retron” cassette encoding a reverse transcriptase, an RNA template (msd), and a primer (msr) that together synthesize a single-stranded, covalently linked RNA-DNA molecule (3). A major advantage of this tool comes from the authors’ finding that the msd sequence can be reprogrammed to include a short variable sequence (about 75 nucleotides) that can target and modify complementary sequences in the genome without disrupting the reverse transcriptase activity of the retron cassette. In an initial experiment, beta recombinase and an engineered retron cassette were coexpressed in bacterial cells to revert deleterious stop codons in a *kanR* resistance gene; the result was increased cell survival when selecting for kanamycin resistance.

Another challenge, however, is to engineer synthetic circuits that operate autonomously in living cells by accurately detecting and responding to input information. Trigger-inducible gene expression systems allow adjustable gene expression in response to specific input levels. For example, in biomedical setups, microencapsulated mam-

malian cells are expected to autonomously take over essential therapeutic tasks in the patient’s body when engineered with closed-loop gene networks. These engineered cells have been shown to control hyperglycemia (8) and urate homeostasis (9) when plugged into the metabolism of mice. Engineered gene circuits are useful not only for therapeutic applications but also in diagnostics to classify cell types (10) or to detect allergic reactions in human blood (11).

The work of Farzadfard and Lu exploited a variety of genetic regulatory circuits to drive the expression of the engineered retron as well as beta recombinase. This system was able to provide an analog recording of the magnitude and duration of a treatment by a chemical inducer. The recording was stored in the distributed genomes of a living bacterial cell population and was recoverable over a long time period. Moreover, it is possible to increase the number of engineered retrons to target diverse genomic regions or reset specific, genomically encoded memory information (see the figure). The three functions “write” (beta recombinase), “input” (msd template), and “read” (reporter enzyme) can also be functionally decoupled by putting them under the control of different gene expression units; this strategy results in a three-input AND gate, which highlights the potential of applying this framework to the design of more complex memory and computational circuits (10, 12).

It is the low recombination efficiency of this gene-editing platform that enables the analog recording of input signals in bacterial subpopulations. This setup could be applied in diagnostic settings by using bacterial cells to generate a long-term recording of a set of environmental events that

can be subsequently analyzed using DNA sequencing methodologies (as shown in this report) to score recombinant frequencies. Substantially enhancing the recombination efficiency of the gene-editing platform would allow the efficient manipulation of genomes on a single-cell level and would also provide a range of other applications, including digital biocomputational devices.

Recently, another report described an inducible gene-editing platform in mammalian cells, based on the CRISPR/Cas9 system (13). In this case, a genomically integrated, inducible Cas9 nuclease was used to introduce DNA strand breaks at specific guide RNA (gRNA)-targeted loci to direct the homologous recombination of an exogenously provided ssDNA template. Retrons have also been shown to be functional in mammalian cells (14). The combination of this CRISPR/Cas9-based platform with programmable control of gRNA expression as well as retron-based ssDNA template production could herald a new era of dynamic genome editing in mammalian cells. Coupled with recent advances in gene circuit design, such a system could provide prosthetic cell implants with the capacity to record and display information about disease states in the genome. It could also be used to dynamically program gene circuits to fine-tune tailored therapeutic interventions.

Despite our limited understanding of retrons and the long-term impact of genome engineering in living cells, the extent to which Farzadfard and Lu could combine naturally derived tools to build a new system with novel functionalities is remarkable. These new synthetic biology-inspired engineering approaches will provide us with strategies for combining gene control modalities with genome-editing technologies and will allow the construction of the next generation of “computerized” living cells. ■

## REFERENCES

1. W. Weber, M. Fussenegger, *Nat. Rev. Genet.* **13**, 21 (2012).
2. P. Mali, K. M. Esvelt, G. M. Church, *Nat. Methods* **10**, 957 (2013).
3. F. Farzadfard, T. K. Lu, *Science* **346**, 1256272 (2014).
4. T. S. Gardner, C. R. Cantor, J. J. Collins, *Nature* **403**, 339 (2000).
5. B. P. Kramer et al., *Nat. Biotechnol.* **22**, 867 (2004).
6. J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, D. Endy, *Science* **340**, 599 (2013).
7. P. Siuti, J. Yazbek, T. K. Lu, *Nat. Biotechnol.* **31**, 448 (2013).
8. D. Ausländer et al., *Mol. Cell* **55**, 397 (2014).
9. C. Kemmer et al., *Nat. Biotechnol.* **28**, 355 (2010).
10. Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, Y. Benenson, *Science* **333**, 1307 (2011).
11. D. Ausländer et al., *Nat. Commun.* **5**, 4408 (2014).
12. S. Ausländer, D. Ausländer, M. Müller, M. Wieland, M. Fussenegger, *Nature* **487**, 123 (2012).
13. F. González et al., *Cell Stem Cell* **15**, 215 (2014).
14. O. Mirochnitchenko, S. Inouye, M. Inouye, *J. Biol. Chem.* **269**, 2380 (1994).

# An economic perspective on the EPA's Clean Power Plan

Cross-state coordination key to cost-effective CO<sub>2</sub> reductions

By Meredith Fowle,<sup>1\*</sup> Lawrence Goulder,<sup>2</sup> Matthew Kotchen,<sup>3</sup> Severin Borenstein,<sup>1</sup> James Bushnell,<sup>4</sup> Lucas Davis,<sup>1</sup> Michael Greenstone,<sup>5</sup> Charles Kolstad,<sup>2</sup> Christopher Knittel,<sup>6</sup> Robert Stavins,<sup>7</sup> Michael Wara,<sup>2</sup> Frank Wolak,<sup>2</sup> Catherine Wolfram<sup>1</sup>

In June, the Obama Administration unveiled its proposal for a Clean Power Plan, which it estimates would reduce carbon dioxide (CO<sub>2</sub>) emissions from existing U.S. power plants 30% below 2005 levels by 2030 (see the chart). Power plant emissions have declined substantially since 2005, so the plan is seeking reductions of about 18% from current levels. Electricity generation accounts for about 40% of U.S. CO<sub>2</sub> emissions.

The Environmental Protection Agency (EPA) is seeking public comment on the proposed plan before issuing a revised and final rule in June 2015. The proposed plan, which is likely to undergo substantial

**POLICY** revision, has been shaped by a confluence of technical, political, legal, and economic factors. Here, we point out, from an economics perspective, some of the main attractions and potential weaknesses of the current proposal. We focus on key design features that will affect whether the plan achieves its intended emissions reductions and on what can be done to help minimize the economic costs of meeting the plan's requirements.

In the absence of new federal legislation, the main channel for a national climate-change mitigation policy is action by the Executive Branch under authority of the Clean Air Act. The plan seeks to regulate emissions from fossil fuel-fired power plants in three steps. First, the EPA identifies "adequately demonstrated" approaches for reducing emissions in the power sector. Second, the EPA derives state-specific emissions standards based on these demonstrated approaches, taking into account the charac-

teristics of each state's power sector. Third, states are obligated to design and implement plans to reduce their emissions and meet the EPA's designated standard.

There are vast uncertainties associated with estimates of the benefits and costs of the plan. The required Regulatory Impact Analysis accompanying the plan (1) estimates that, under most scenarios, benefits will exceed costs by a considerable margin. One central estimate puts benefits in the year 2030 at \$76 billion, with projected compliance costs of \$9 billion. Estimated benefits include avoided climate-related damages in both the United States and other countries. The Administration's rationale for including other countries is that addressing global climate change will require all countries to pursue policies based on consideration of global costs and benefits associated with their own actions.

Also included in the benefits are domestic "nonclimate" effects. To the extent that the plan leads to reductions in CO<sub>2</sub> emissions from fossil-fuel plants, there will also be reductions in other pollutants, with associated benefits to health and the local environment. Drawing from existing studies, the EPA considered health benefits from reductions in particulate matter and ground-level ozone. These reductions were estimated to yield about 60% of the plan's gross benefits.

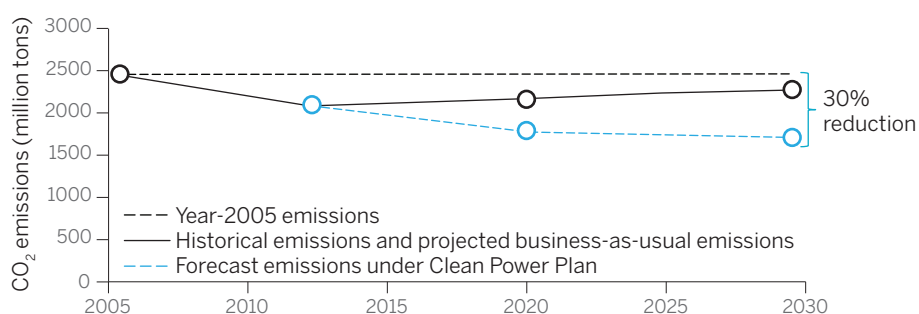
**WILL INTENDED REDUCTIONS BE ACHIEVED?** Although the Administration hopes that the plan will bring about emissions reductions of a given magnitude, the plan does not specify particular levels

of emissions reductions that states must achieve. Instead, compliance requirements are expressed as an emissions ratio, building upon EPA's considerable experience with regulatory requirements defined in this way. The numerator of the ratio measures pounds of CO<sub>2</sub> emitted by sources covered under the plan. The denominator is a measure of electricity generation at most existing sources, with an adjustment to account for generation avoided due to demand-side efficiency improvements. The quantity of emissions associated with meeting the standards depends on both the level of emissions (in the numerator) and the level of electricity generation (in the denominator). The Administration's estimate of emissions reductions thus depends on its forecasts of electricity generation.

Each state has the option to convert the ratio-based requirement into a mass-based limit on the quantity (i.e., tons) of CO<sub>2</sub> emitted. In principle, this involves multiplying projected electricity production at affected generating units in the state by the target emissions ratio. In practice, the modeling and assumptions required to make the state-level conversions will be complicated, and many details are unspecified in the proposed plan.

The use of a ratio-based standard makes for a loose connection between meeting the required emissions-output ratio and achieving the Administration's forecast emissions reductions. If a state chooses to maintain a ratio-based standard rather than convert to a mass-based limit, it could potentially meet the EPA's requirement without lowering emissions. A state could bring down its average emission ratio by simply increasing production of low-carbon—but not necessarily no-carbon—electricity. Although the extent to which this will undermine the plan's goal for emission reductions is uncertain, economic analysis of ratio-based standards in other sectors has revealed considerable potential for these unintended effects (2).

## Projected CO<sub>2</sub> emissions impacts of the proposed power plan



Data from (1).

<sup>1</sup>University of California Berkeley, Berkeley, CA 94720, USA. <sup>2</sup>Stanford University, Palo Alto, CA 94305, USA. <sup>3</sup>Yale University, New Haven, CT 06511, USA. <sup>4</sup>University of California Davis, Davis, CA 95616, USA. <sup>5</sup>University of Chicago, Chicago, IL 60637, USA. <sup>6</sup>Massachusetts Institute of Technology, Cambridge, MA 02412, USA. <sup>7</sup>Harvard University, Cambridge, MA 02138, USA. \*E-mail: fowle@berkeley.edu



Another issue with ratio-based standards concerns programs for demand-side energy efficiency improvements (such as programs designed to increase the adoption of more-efficient appliances). These can offer a cost-effective way for states to cut emissions by reducing electricity consumption. To receive credit for energy efficiency improvements under a ratio-based standard, efficiency-induced reductions in electricity demand are added to the denominator in calculating the state's emissions ratio. The problem is that accurately measuring how efficiency investments affect electricity demand is notoriously difficult, and recent studies indicate that savings are frequently overestimated by a wide margin (3). Overestimation of gains from energy efficiency will in effect reduce the stringency of a ratio-based standard.

Under a mass-based approach, the two concerns discussed above no longer apply. There is no incentive for a state to increase electricity production so as to reduce an emissions ratio. Efficiency measures would contribute directly to compliance under a cap by reducing demand for electricity, and therefore emissions, without any need to estimate savings for compliance purposes. A mass-based approach would also be administratively simpler and would provide greater certainty about the plan's impact on emissions. Moreover, for states that already have mass-based policies limiting greenhouse gas emissions, mass-based standards are more easily integrated into the existing regulatory framework.

Despite the various attractions of a mass-based approach, it is far from certain that states will choose this compliance option. Ratio-based standards have more built-in flexibility. While maintaining the same emissions ratio, states can expand and contract emissions in response to changing economic conditions that affect electricity demand and supply.

To ensure that ratio-based standards meet their objective, the EPA should establish detailed protocols and guidance for evaluation, monitoring, and verification of efficiency programs. The EPA should favor field-based savings estimates over engineering estimates and should encourage evaluators to take advantage of state-of-the-art approaches to program evaluation.

**WILL WE PAY TOO MUCH FOR EMISSIONS REDUCTIONS?** To reduce emissions cost-effectively, it is critical to adopt a flexible structure in which lower-cost options can be favored over higher-cost ones. Several features of the plan are consistent with this flexibility principle.

The plan offers within-state flexibility in how each state can comply with its standard. When setting the standard for each state, the EPA considered a range of demonstrated methods for reducing emissions, including efficiency improvements at power plants, greater use of natural-gas plants, expansion of renewables, and demand-side energy efficiency programs. Yet states need not employ these approaches. They can choose any approach, taking advantage of changing market conditions, new information, and cost-saving innovations to comply at least cost.

The plan also allows for across-state flexibility in terms of where required emissions reductions occur. Across-state coordination has important cost implications because low-cost abatement options are not evenly distributed across states. Emissions trading is one approach to coordinating emissions abatement cost-effectively across states. Under the proposed plan, states can

### ***"a state ... could potentially meet the EPA's requirement without lowering emissions"***

file multistate implementation plans that allow electricity producers in one state to trade emissions reduction obligations with producers in other states. Plants facing relatively high costs of emission reductions could pay plants with relatively low costs to take on additional emissions reductions on their behalf. Experience in the United States with emissions trading systems has demonstrated that this kind of flexibility within and across states can substantially lower the costs of reducing emissions overall (4).

Even if states do not pursue emissions trading, effective interstate coordination on implementation within regional electricity markets will be critical. Electricity production is linked and dispatched through regional grid interconnections that span multiple states. Electricity will be traded among states within a regional power system regardless of whether the states choose to trade emissions. Without interstate coordination, technologically identical power plants located in different states—yet servicing the same electricity customers—could face very different regulatory incentives and operating costs. Such differences would distort the flow of electricity between states and raise the overall cost per ton of emissions reduced under the plan.

At a minimum, states within the same electricity interconnection should coordinate to implement the plan in a way that

harmonizes emissions reduction incentives across states. Although coordination of state implementation is certainly possible under the proposed plan, it is by no means guaranteed. The EPA can promote coordination by issuing specific guidelines—"model rules"—in the revised plan that states seeking to coordinate can adopt. The EPA has assumed this coordinating role in past programs to reduce administrative difficulties that impede coordination across state boundaries. Under the proposed plan, the EPA will also need to consider how coordination among states can be achieved between rate-based and mass-based approaches to compliance.

**KEY ISSUES GOING FORWARD.** The Clean Power Plan is the centerpiece of the Obama Administration's Climate Action Plan. The proposed reductions in power sector emissions may take on strategic importance in an international setting, where other countries are looking to the United States to gauge their own commitments to reduce emissions as part of the next international agreement.

Two concerns need to be addressed regarding the ratio-based approach to defining emissions reduction targets. The first is perverse incentives for expanded electricity production in place of reduced emissions. The second is potential overestimation of energy efficiency gains that will effectively weaken the standard. The plan would also benefit from EPA guidance as to how states will be permitted to convert ratio-based targets into mass-based approaches.

The plan's considerable flexibility regarding how and where emission reductions can occur is an important feature because it promotes cost-effectiveness. Whether states will fully capitalize on this flexibility is an open question. Effective coordination among states and within multistate power markets will require state governments to overcome considerable analytical, logistical, and administrative hurdles. As part of the final rule, the EPA can play a facilitating role by providing additional and more concrete guidance on how states choosing to harmonize their implementation plans and/or compliance obligations can do so efficiently. These considerations are critical to meeting emissions reduction goals at the least cost. ■

#### REFERENCES

1. U.S. EPA, *Regulatory Impact Analysis for the Proposed Carbon Pollution Guidelines for Existing Power Plants and Emission Standards for Modified and Reconstructed Power Plants* (EPA, Washington, DC, 2014).
2. S. P. Holland, J. E. Hughes, C. R. Knittel, *Am. Econ. J.: Econ. Policy* **1**, 106 (2009).
3. H. Allcott, M. Greenstone, *J. Econ. Perspect.* **26**, 3 (2012).
4. R. Schmalensee, R. N. Stavins, *J. Econ. Perspect.* **27**, 103 (2013).

10.1126/science.1261349

## PUBLIC HEALTH

# Malaria: Declassified

By Irwin W. Sherman

**M**alaria has plagued humankind for millennia. During World War II, as American and Allied forces engaged in battles in North Africa, Asia, and the Pacific, troop losses due to malaria were sometimes as great as those due to bullets and bombs. With the Japanese occupation of Java (today Indonesia), the Western forces found themselves abruptly cut off from a major supplier of the only known curative agent—quinine, a compound derived from the bark of the cinchona tree. An urgent search for other antimalarials to protect the troops was soon under way. These efforts, led by the U.S. War Department, are the subject of *The Malaria Project: The U.S. Government's Secret Mission to Find a Miracle Cure* by science journalist Karen M. Masterson.

The impetus for this book came when Masterson discovered declassified letters, reports, and data sheets on the antimalarial program in the U.S. National Archives. She then conducted extensive interviews with the few remaining witnesses, their kin, and a number of scientists and public health experts, including historian of medicine Leo Slater and malaria expert William Collins of the U.S. Centers for Disease Control and Prevention. The result is a fascinating story of the scientists who committed themselves to curing malaria.

After a succinct introduction to the disease itself, Masterson's lucid writing takes the reader on a journey to 1920s Germany. It was here, in the laboratories of the pharmaceutical company Bayer, that imaginative chemists first sought to develop a "magic bullet" cure for malaria.

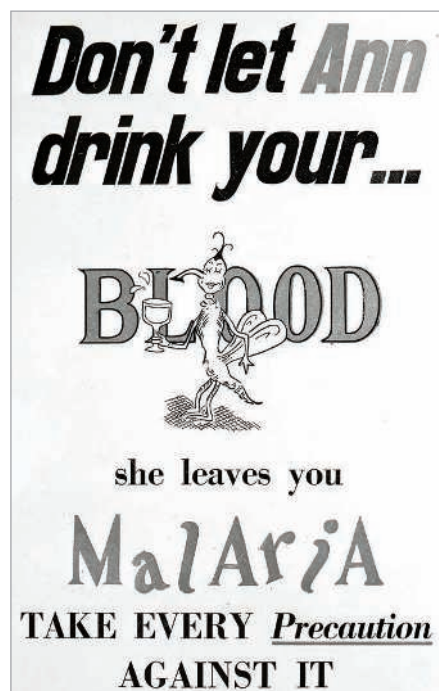
In 1939, top malaria experts, renal physiologists, organic chemists, and parasitologists were recruited to a medical "Manhattan Project" focused on eradicating malaria in the United States. The project was top secret, and the conceptual framework was ideal for research—a bottom-up approach in which scientists steered the course. By 1944, more than 400 investigators were working on drug synthesis, animal testing, and human studies at institutions across the country un-

**The Malaria Project**  
The U.S. Government's Secret Mission to Find a Miracle Cure  
Karen M. Masterson  
New American Library, 2014.  
432 pp.



der the supervision of the Board for Coordination of Malarial Studies. In the end, more than 14,000 compounds were cataloged, two of which would go on to play major roles in worldwide campaigns to eradicate malaria.

In addition to providing an in-depth history of malaria research, Masterson allows readers to become fully immersed in the military campaigns in Africa, the Solomon Islands, and the Philippines, where malaria was rampant. She recounts how supply and equipment problems faced by medically ignorant military leaders like George Patton were trumped by the dangers that ponds and puddles (breeding grounds for malaria-carrying mosquitos) posed to troops. Flagrant disregard of the recommendations of scientists and the medical community translated into untreated fevers on the front line.



An antimalaria poster, designed by Captain Theodor S. Geisel, later known as Dr. Seuss.

In an effort to encourage compliance within the armed forces, the War Department commissioned posters and films designed to coerce, manipulate, and entertain the troops into following approved protocols for malaria prevention and control. The cartoonist, Theodor S. Geisel (Dr. Seuss), then an Army captain, created comics that warned of deadly Ann, a female mosquito that drinks G.I. blood. Cartoon trailers produced by Warner Brothers dramatized how malaria spread, and Walt Disney released a short film in which the seven dwarves learned about the menace of malaria.

During a campaign in Tunis in early 1944, Allied forces recovered an antimalarial drug called sontochin, originally developed by Bayer chemists. Sontochin was found to be identical to one of the compounds developed by the U.S. researchers—SN-183. Although it could cure ducklings infected with a bird malaria, SN-183 had yet to be tested in humans. Human trials were subsequently conducted on inmates at Stateville Prison in Illinois. Later renamed chloroquine, SN-183 went on to become one of the mainstays of the World Health Organization's campaign to eradicate malaria in the 1950s. Also tested at Stateville was compound SN-13,272, known later as primaquine, which is the only drug to date that can eradicate the liver stages of malaria and eliminate relapse of *Plasmodium vivax* malaria, the most widespread of human malarias.

Interestingly, Andrew C. Ivy, a well-known and respected physician who testified on behalf of the American Medical Association at the Nuremberg Medical Trial for Nazi war criminals in 1946, actually approved of the trials conducted at Stateville. Masterson quotes from his assessment, which states that the malaria drug trials were "ethically acceptable," "ideal," and conformed "with the highest standards of human experiments." Today, as Masterson correctly observes, these trials would probably not pass muster.

The book is not without errors. For example, Masterson states that "some scientists believe [the] parasite originated as a bird malaria"; however, it is generally agreed that the human parasites originated in primates. Also, the genus of all malarias is *Plasmodium*, not *Plasmodia*. Despite these and a few other small issues, *The Malaria Project* is a fascinating and well-told account of the search for a cure for malaria. Its strength lies in emphasizing a lesson not to be forgotten: that drug and vaccine development alone will not eliminate malaria. Eradication requires education, improved health care, mosquito control, and poverty intervention.

10.1126/science.1261450



## BIG DATA

# The big data debate

By Fred H. Cate

**B**ig data—the collection, aggregation or federation, and analysis of vast amounts of increasingly granular data—present serious challenges not only to personal privacy but also to the tools we use to protect it. *Privacy, Big Data, and the Public Good* focuses valuable attention on two of these tools: notice and consent, and de-identification—the process of preventing a person's identity from being linked to specific data. The book presents a collection of essays from a variety of perspectives, in chapters by some of the heavy hitters in the privacy debate, who make a convincing case that the current framework for dealing with consumer privacy does not adequately address issues posed by big data.

In 1998, the U.S. Federal Trade Commission identified notice and consent as the “most fundamental” principles for protecting privacy (1). In 2012, the U.S. Consumer Privacy Bill of Rights included as its first principle: “Consumers have a right to exercise control over what personal data companies collect from them and how they use it” (2). The pending draft of the European Union Data Protection Regulation refers to “consent” more than 100 times (3). Regulators, it seems, are entranced by notice and consent. It is perhaps not surprising, then, that we are inundated with privacy notices and demands that we click “I agree” when we visit banks, doctors, Web sites, app stores, and the like. Yet, as Solon Barocas and Helen Nissenbaum point out in Chapter 2, there is “overwhelming evidence that most of us neither read nor understand them.”

The challenge presented by big data lies not just in the size of the data sets or the ubiquity with which data are collected but also, as Paul Ohm points out in Chapter 4, in the fact that big data “thrives on surprising correlations and produces inferences and predictions that defy human understanding.” One can readily agree when he asks, “How can you provide notice about the unpredictable and unexplainable?”

In a May 2014 report (4), the U.S. President's Council of Advisors on Science and Technology described the framework of no-

tice and consent as “unworkable as a useful foundation for policy.” The report goes on to add that “only in some fantasy world do users actually read these notices and understand their implications before clicking to indicate their consent.”

The authors of *Privacy, Big Data, and the Public Good* are generally unwilling to go so far. Despite pages of criticism, they advocate for tweaking the current system so that notice and choice are used more effectively. In Chapter 2, for example, Barocas and Nissenbaum propose that notice and consent be required in cases when the proposed use of the data departs from the context in which the data were collected. Alternatively, in



Chapter 4, Ohm proposes that groups similar to institutional review boards might develop ethical norms that would determine when consent was necessary.

The book's critique of de-identification is equally compelling. Currently, de-identified data are exempt from most privacy laws. Yet throughout the book, the authors point to a number of well-publicized examples in which de-identified data sets have been released, only to have the data re-identified within days. In short, the more data are available, the harder it is to de-identify them effectively. Or, as Cynthia Dwork succinctly sums up the situation in Chapter 14: “De-identified data’ isn’t.”

On this subject, the authors offer an array of responses, ranging from Dwork's “differential privacy” approach, which uses statistical analyses to assess privacy risks presented by repeated queries of the same data set, to Ohm's recommendation that de-identified data should no longer be exempt from privacy laws.

## Privacy, Big Data, and the Public Good

Frameworks for Engagement

Julia Lane, Victoria Stodden,

Stefan Bender, and

Helen Nissenbaum, Eds.

Cambridge University Press, 2014.

342 pp.



There is a third aspect of the book that warrants special mention—a series of chapters addressing the uses of big data in the context of cities, urban planning, and research. These areas are not often discussed in the context of big data, and although they make the book feel a bit eclectic, they address new and interesting issues that clearly merit attention.

As society becomes more “datafied” (5)—a term coined to describe the digital

quantification of our existence—our privacy is ever more at risk, especially if we continue to rely on the tools that we employ today to protect it. *Privacy, Big Data, and the Public Good* represents a useful and approachable introduction to these important issues.

## REFERENCES AND NOTES

1. U.S. Federal Trade Commission, *Privacy Online: A Report to Congress* (FTC, Washington, DC, 1998).
2. The White House, *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation* (White House, Washington, DC, 2012).
3. *Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation)*, Unofficial Consolidated Draft Text, 22 October 2013.
4. President's Council of Advisors on Science and Technology, *Big Data and Privacy: A Technological Perspective* (PCAST, Washington, DC, 2014).
5. V. Mayer-Schönberger, K. Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (Houghton Mifflin, Boston, 2013).

The reviewer is at the Indiana University Maurer School of Law, Bloomington, IN 47405, USA. E-mail: fred@fredhcate.org

## LETTERS



Young fishers in Burkina Faso.

Edited by Jennifer Sills

### ***Fauna in decline: A big leap to slavery***

THE POLICY FORUM by J. S. Brashares *et al.* ("Wildlife decline and social conflict," 25 July, p. 376) links environmental degradation to child slavery and human trafficking. However, the causal pathways between wildlife decline and social outcomes, especially child slavery, have not been empirically tested; it is likely that wildlife declines are only a tiny component of the causes of exploitative child labor and slavery.

While the linkages Brashares *et al.* diagram are all plausible, and some have abundant empirical support, we question the strength of association between the exploitation of global fisheries and child slavery. To claim or imply that resolving the global fisheries commons problem will mitigate such despicable social practices is dangerous.

Surtees (1) lists economic, political, legal, social, and individual characteristics as preconditions for child trafficking for forced labor and sexual exploitation. Noticeably absent in her report is any mention of a link between wildlife decline, or any form of environmental degradation, as a precondition for child trafficking.

Conservation goals have increasingly been tied to aspects of human well-being and social outcomes, but it is important not to overestimate the negative impacts of environmental degradation when they are likely to play a secondary role in driving social outcomes. Doing so risks exacerbating the very social problems they purportedly aim to tackle. There is no doubt that greater stresses on the environment put greater economic strains on families

that rely heavily on the environment for their livelihoods, yet the evidence the authors cite is unsatisfactory and risks moving resources to combat child labor and slavery toward less effective programs. More effective ways to combat the exploitation of children could include better enforcing child labor laws, increasing educational opportunities and services, and increasing poverty reduction and family planning programs.

The scientific community has a responsibility to not overstate the potential impacts of the environment on social outcomes for another reason as well. Policy in democratic countries, and especially environmental policy, depends on public support. The public is increasingly wary of the political framing or spinning of issues by conflating environmental issues with other concerns, regardless of support.

**Yuta J. Masuda<sup>1</sup>\* and Tim Scharks<sup>2</sup>**

<sup>1</sup>Central Science, the Nature Conservancy, Boston, MA 02111, USA. <sup>2</sup>Evans School of Public Affairs, The University of Washington, Seattle, WA 98195, USA.

\*Corresponding author. E-mail: ymasuda@tnc.org

#### REFERENCE

1. R. Surtees, "Child trafficking in Sierra Leone" (UNICEF, 2005).

### ***Fauna in decline: Management risks***

IN THEIR POLICY FORUM "Wildlife decline and social conflict" (25 July, p. 376), J. S. Brashares *et al.* claim that a decline in the resource base will increase exploitation effort and slave labor. Using fisheries to illustrate this claim, the authors conjecture that a stock decline increases effort, which further degrades the stock. However,

the contention that decreased stocks will increase effort is sharply contradicted by theoretical and empirical literature on open-access fisheries (1–3), as well as modeling of collapse in pre-modern wildlife-dependent societies (4). Although evidence from fisheries is limited, it appears that when stock declines, fishing effort declines as well (2, 3, 5–8).

Potential links between social conflict and resource degradation are not implausible. Globalization increases the potential value of natural resources in the absence of degradation. Unfortunately, without effective governance, globalization can also increase incentives for overexploitation and illicit activity (9, 10). Brashares *et al.* rightly note that policies should look at underlying causes, and I agree with their endorsement of resource tenure (rules that govern access to resources) and international commitments to user rights for fishers. However, real-world resource management often falls short of providing user rights and full institutional reform. When fisheries are managed to stabilize or increase stocks without restricting access to the resource base, effort actually increases (6–8). If Brashares *et al.* are correct that lower stocks equate to higher effort, policies that improve biological stocks could indeed decrease effort and associated slave labor. However, if they are wrong, as the literature suggests, biological management alone could worsen labor outcomes.

**Martin D. Smith**

Nicholas School of the Environment and Department of Economics, Duke University, Durham, NC 27708, USA. E-mail: marsmith@duke.edu

#### REFERENCES

1. H. S. Gordon, *J. Polit. Econ.* **75**, 124 (1954).
2. T. Bjørndal, J. M. Conrad, *Can. J. Econ.* **20**, 74 (1987).
3. C. Liese *et al.*, *Environ. Dev. Econ.* **12**, 123 (2007).
4. V. L. Smith, *J. Polit. Econ.* **83**, 727 (1975).
5. J. E. Cinner, T. Daw, T. R. McClanahan, *Conserv. Biol.* **23**, 124 (2009).
6. R. Cudney-Bueno, X. Basurto, *PLOS One* **4**, e6253 (2009).
7. F. R. Homans, J. E. Wilen, *J. Env. Econ. Manage.* **32**, 1 (1997).
8. M. D. Smith *et al.*, *J. Env. Econ. Manage.* **55**, 265 (2008).
9. M. D. Smith *et al.*, *Science* **327**, 784 (2010).
10. J. E. Wilen, *Mar. Res. Econ.* **28**, 203 (2013).

#### ***Response***

THE DISCOMFORT EXPRESSED by Masuda and Scharks in linking wildlife declines to social conflict and Smith's concern about harvest effort increasing when wildlife declines reinforce the central message of our Policy Forum: Despite growing evidence of the importance of wildlife-society linkages, these connections are rarely incorporated in the work of researchers, policy-makers, or practitioners. We did not aim to test causal relationships or economic theories linking wildlife declines to child



slavery. Our primary goal was to illustrate the deep environmental roots of many types of social conflict and elevate the visibility of this oft-ignored driver of social unrest.

As we acknowledged in our Policy Forum, wildlife decline is only one of a number of drivers of social conflict. Yet, there is growing evidence that wildlife decline is in many cases far more than a “tiny component” of various forms of social injustice, including child labor. The fact that the fishing industry is a primary user of child and forced labor in many parts of the world has been well documented (1–5). The United Nations directly relates fish stock declines to exploitative labor practices, noting that decreasing fish stocks increases labor demand and makes workers more “vulnerable to human trafficking at sea” (1). UNICEF includes the loss of fishing livelihoods as a driver of child trafficking (5).

Millions of people rely directly on fisheries and terrestrial wildlife for their health, well-being, and livelihoods; to understand the ways in which wildlife decline is exacerbating labor exploitation and other social ills, we must address these links. Although fields such as political ecology have built a body of evidence over the past 30 years elaborating such connections, wildlife and society are still too often treated as separate spheres.

We do not call for policies that divert resources from other efforts to address social conflicts, as feared by Masuda and Scharks. Instead, we hope that by considering underlying drivers of social ills, including wildlife decline, policies and programs will extend beyond superficial treatment of proximate outcomes. Failing to recognize connections between wildlife decline and social injustice in the face of existing evidence will at best waste resources on inadequate efforts, and at worst exacerbate these problems through unintended consequences.

Smith argues that when populations of harvested wildlife decline, the effort expended by fishermen and hunters to collect these resources will decline. We question the logic of and evidence for this proposition at both the individual and aggregate levels. Evidence for individuals increasing effort in the face of wildlife decline is abundant in the literature and applies to harvests of whales, reef fish, cod, and terrestrial mammals (6–9). Theories from applied economics and governance reinforce this argument: The labor debt created by the increased effort required to harvest a declining resource often contributes to exploitative labor practices (10).

At the aggregate level, harvesters in some systems may stop hunting or fishing if resources decline below a threshold level, but there is overwhelming evidence that overall effort can, and has, increased in many cases where stocks were diminished (11). Studies have shown that individuals who lack access to alternative livelihoods are unlikely to exit a declining fishery (6), and global fishing effort has grown 10-fold since 1950, while total catch has declined (12). As Donald Kennedy recently summarized, “Declines in the fish population encourage...larger, longer-range, and higher-capacity vessels. That further intensifies the pressure on the resource, so the process has its own positive feedback” (13).

We agree with Smith that increases in harvest effort following wildlife declines will be stronger in contexts characterized by “poverty and institutional failure.” Such situations represent the norm in many regions of a world where nearly 3 billion people live in situations of poverty and weak governance (14, 15).

The interplay between wildlife decline, harvest effort, and the demand for slave labor is an underappreciated social-environmental problem for which the drivers and outcomes are poorly understood. While we disagree with Smith about how harvest effort commonly adjusts to wildlife scarcity, we agree that gaining a better understanding of the complexities of these relationships is a critical step in addressing social conflict related to wildlife decline. We also emphasize that many of the outcomes we outlined in our Policy Forum (such as political stability and regional violence) operate independently of the labor issue that concerns Smith.

**Lauren Withey,<sup>1</sup> Katherine Seto,<sup>1</sup> Douglas J. McCauley,<sup>2</sup> Kathryn J. Fiorella,<sup>1</sup> Ryan A. Marsh,<sup>1</sup> Briana Abrahms,<sup>1</sup> Tristan A. Nuñez,<sup>1</sup> Christopher D. Golden,<sup>3</sup> Justin S. Brashares<sup>1\*</sup>**

<sup>1</sup>ESPM, University of California, Berkeley, CA 94720, USA. <sup>2</sup>EEMB, University of California, Santa Barbara, CA 93106, USA. <sup>3</sup>HEAL Program, Wildlife Conservation Society, Bronx, NY 10460, USA.

\*Corresponding author. E-mail: brashares@berkeley.edu

#### REFERENCES

1. U.N. Office on Drugs and Crime, “Transnational organized crime in the fishing industry” (UNDOC, Vienna, 2011).
2. U.S. Department of Labor, “Findings on the Worst Forms of Child Labor” (2012).
3. EJF, “All at sea: The abuse of human rights aboard illegal fishing vessels” (Environmental Justice Foundation, London, 2010).
4. FishWise, “Trafficked II: An updated summary of human rights abuses in the seafood industry” (FishWise, Santa Cruz, CA, 2014).
5. UNICEF, “Training manual to fight trafficking in children for labour, sexual, and other forms of exploitation” (UNICEF International Labour Office, Geneva, 2009).
6. J. E. Cinner et al., *Glob. Environ. Change* **21**, 7 (2011).

7. C. W. Clark, R. Lamberson, *Mar. Policy* **6**, 103 (1982).
8. S. Mainka, M. Trivedi, “Links between biodiversity conservation, livelihoods and food security” (IUCN, Gland, Switzerland, 2002).
9. J. Hutchings, R. Myers, in *The North Atlantic Fisheries*, D. Newell, R. Ommer, Eds. (Island Institute Studies, University of Prince Edward Island, Charlottetown, PEI, 1995), pp. 37–94.
10. K. Bales, *Understanding Global Slavery* (University of California Press, Oakland, CA, 2005).
11. R. H. Thurstan, S. Brockington, C. M. Roberts, *Nat. Commun.* **1**, 15 (2010).
12. R. A. Watson et al., *Fish Fish.* **14**, 493 (2013).
13. D. Kennedy, *Science Magazine’s State of the Planet 2006–2007* (Island Press, Washington, DC, 2006).
14. D. Kaufmann, A. Kraay, M. Mastruzzi, “Governance matters VIII” (Social Science Research Network, Rochester, NY, 2009).
15. UNDP, *Reducing Vulnerabilities and Building Resilience* (U.N. Development Program, New York, 2014).

## Fauna in decline: Beyond extinction

IN THEIR REVIEW “Defaunation in the Anthropocene” (25 July, p. 401), R. Dirzo et al. confuse two separate processes that are important to conservation biology, extinction, and population decline. Traditionally, when species are referred to as extinct, it means they no longer have living members. The IUCN made a modification when it recognized “extinct in the wild” to be applied to species that existed only in captivity. Biodiversity loss or gain is usually quantified in terms of species numbers. Unlike extinction, in the case of a population collapse (a 90% decline), there can still be hope for recovery.

Most extinctions have occurred on oceanic islands or in restricted freshwater locations, with very few occurring on Earth’s continents or in the oceans. The world’s greatest conservation problem is not species extinction, but rather the precarious state of thousands of populations that are the remnants of once widespread and productive species.

**John C. Briggs**

Department of Fisheries and Wildlife, Oregon State University, Corvallis, OR 97330, USA. E-mail: clingfishes@yahoo.com



The Wyoming toad (*Bufo baxteri*) has been categorized as extinct in the wild since 1991.

PHOTO: SARA ARMSTRONG/WIKIMEDIA COMMONS

## Fauna in decline: The community way

CURRENT CONSERVATION STRATEGIES to mitigate the impact of climate change on terrestrial biodiversity rely heavily on capture, transfer, and release of single species (single-species translocation), despite the fact that ecological interactions between species are likely to be the first component of the ecosystem to be impacted by climate change (1) before any population or species goes extinct. In their Review ("Reversing defaunation: Restoring species in a changing world," 25 July, p. 406), P. J. Seddon *et al.* analyzed conservation translocations and emphasized the need for "more intensive forms of threatened species management." To conserve functioning ecosystems, management tools should focus on conserving whole communities rather than single charismatic species.

Ecosystem-scale translocation is one way to accomplish this goal: Aboveground and belowground elements of a functioning terrestrial ecosystem (including vegetation and topsoil) are carefully collected and moved together. Small-scale examples of ecosystem-scale translocation have been applied for 30 years for the purpose of ecological restoration under the name of habitat translocation or vegetation direct transfer (2–5). The strategy has proven successful in conserving plant, invertebrate, and microbial communities as well as ecosystem functions (5–9). By moving subsets of ecosystems from climatically unstable regions to more stable ones (10), ecosystem-scale translocation provides an opportunity to conserve mature and complex ecosystems threatened by climate change.

**Stéphane Boyer**

Department of Ecology, Bio-Protection Research Centre, Lincoln University, Lincoln, 7647, New Zealand. E-mail: stephane.boyer@lincoln.ac.nz

### REFERENCES

1. J. M. Tylianakis, R. K. Didham, J. Bascompte, D. A. Wardle, *Ecol. Lett.* **11**, 1351 (2008).
2. W. D. Cullen, C. P. Wheeler, in *Proceedings of International Land Reclamation and Mine Drainage Conference and Third International Conference on the Abatement of Acidic, Volume 3 of 4: Reclamation and Revegetation Drainage* (U.S. Department of the Interior Bureau of Mines Special Publication SP 06C-94, Pittsburgh, PA, 1994), pp. 83–92.
3. K. Winterhalder, in *Proceedings of Sudbury '95—Mining and the Environment*, T. P. Hynes, M. C. Blanchette, Eds. (Sudbury, ON, 1995), pp. 1189–1198.
4. J. M. Bullock, *Biol. Conserv.* **83**, 199 (1998).
5. C. Ross *et al.*, in *New Zealand Minerals & Mining Conference Proceedings* (2000), pp. 29–31.
6. D. Rodgers, R. Bartlett, R. Simcock, S. Wratten, S. Boyer, in *Proceedings of the 2011 Workshop on Australian Mine Rehabilitation*, O. Nichols, N. Vikuckis, Eds. (JKTech Pty Ltd., Indooroopilly QLD, 2011), pp. 285–303.
7. B. R. Waterhouse, K. L. Adair, S. Boyer, S. D. Wratten, *Basic*



Boreal forest in Yukon, Canada.

- Appl. Ecol.* **15**, 599 (2014).
8. S. Boyer, S. Wratten, M. Pizey, P. Weber, *Pedobiologia (Jena)* **54S**, S99 (2011).
  9. C. H. Watts, B. R. Clarkson, R. K. Didham, *Biol. Conserv.* **141**, 568 (2008).
  10. J. E. M. Watson, T. Iwamura, N. Butt, *Nat. Clim. Chang.* **3**, 989 (2013).

## Fauna in decline: Protect forests now

THE SPECIAL SECTION on Vanishing Fauna (25 July, p. 392) did an excellent job of highlighting the continuing global biodiversity erosion crisis ("Defaunation in the Anthropocene," R. Dirzo *et al.*, Reviews, p. 401) and progress in combating the crisis by using restoration techniques and translocation of animals ("Reversing defaunation: Restoring species in a changing world," P. J. Seddon *et al.*, Reviews, p. 406). However, the most effective, resource-efficient, and safe opportunities for slowing defaunation continue to be overlooked. We must maintain the world's last large remaining areas of intact habitat, including primary forest areas that remain mostly free of large-scale human development.

These primary forest areas have recently been mapped (1, 2) and include the boreal forest regions of Canada, the United States, and Russia, and the tropical forests of the Amazon Basin, Congo Basin, and parts of Indonesia and Papua New Guinea. Canada's boreal forest accounts for approximately 25% of the world's remaining intact primary forest (1). Initiatives in Canada by provincial, indigenous, and federal government entities as well as industry have rapidly pushed forward the levels of protection, successfully setting aside about 450,000 km<sup>2</sup> of primary forest for conservation purposes over the past 15 years (3). Perhaps the greatest leaders in this effort have been indigenous communities and governments, who have developed

comprehensive land-use plans and new management models for their ancestral lands that balance conservation and development (4).

Protecting the world's still-intact ecosystems, including primary forest, avoids the necessity for costly, difficult, and untested approaches to restore and translocate species and environments (5). All nations should be focused on protecting intact ecosystems now rather than hoping to restore them later.

**Jeffrey V. Wells**

International Boreal Conservation Campaign, Seattle, WA 98101, USA. E-mail: jeffwells@intlboroal.org

### REFERENCES

1. B. Mackey *et al.*, *Conserv. Lett.* 10.1111/conl.12120 (2014).
2. M. C. Hansen *et al.*, *Science* **342**, 850 (2013).
3. J. Wells, *Stud. Avian Biol.* **41**, 1 (2011).
4. P. Badiou *et al.*, "Conserving the world's last great forest is possible: Here's how" (International Boreal Conservation Science Panel, 2013); <http://borealscience.org/wp-content/uploads/2013/07/conserving-last-great-forests1.pdf>.
5. A. Ricciardi, D. Simberloff, *Science* **345**, 884 (2014).

### ERRATA

**Erratum for the Research Article: "mTOR- and HIF-1 $\alpha$ -mediated aerobic glycolysis as metabolic basis for trained immunity" by S.-C. Cheng *et al.*, *Science* **346**, aaa1503 (2014).** Published online 7 November 2014; 10.1126/science.aaa1503

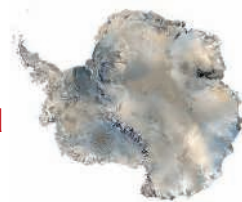
**Erratum for the Perspective: "Clogging information flow in ALS" by J. W. Paul III and A. D. Gitler, *Science* **346**, 1261739 (2014).** Published online 10 October 2014; 10.1126/science.1261739

**Erratum for the Report: "Observation of the transition state for pressure-induced BO<sub>3</sub>→BO<sub>4</sub> conversion in glass" by T. Edwards *et al.*, *Science* **345**, 1261201 (2014).** Published online 19 September 2014; 10.1126/science.1261201



# RESEARCH

Ancient Antarctic glaciers were more dynamic than anticipated  
Woodard et al., p. 847



## IN SCIENCE JOURNALS

Edited by Stella Hurtley

### SEXUAL CONFLICT

#### Why some male animals kill infants

One of the most unpleasant aspects of social life in some animal species is killing of the young by adult males. Lukas and Huchard looked at mammalian groups with a variety of social systems—from mice to mongoose and from bats to bears. Infanticidal behavior in males appeared to be a result of sexual conflict in social species with nonseasonal breeding. Killing of a female's young by a new male will speed her return to a reproductive state and allow him to raise his own young. Evolutionarily, the only successful defense in females appears to be polyandry: Females that mate with multiple males make it hard for any one male to know that he is, or is not, the father of her offspring. — SNV

*Science*, this issue p. 841



### VIRAL INFECTION

#### Flagellin gives rotavirus a one-two punch

Rotavirus causes gastroenteritis, which can be especially severe in infants and young children. The bacterial protein flagellin activates the innate immune system and protects mice against a variety of infectious and inflammatory agents. Zhang *et al.* now report that flagellin both prevented rotavirus infection in mice and cured mice chronically infected with rotavirus. It did so by activating two distinct innate immune signaling pathways, which led to cells in the infected mice producing the cytokines interleukin 22 and interleukin 18.

Similar to flagellin, treating mice with both of these cytokines prevented or cured rotavirus infection. — KLM

*Science*, this issue p. 861

### CLIMATE CHANGE

#### Striking when hot, and more when hotter

Lightning occurs more frequently when it is hotter than when it is colder, but how much more lightning should we expect as global temperatures increase? Currently there are around 25 million lightning strikes per year. Romps *et al.* constructed a proxy based on the energy available to make air

rise in the atmosphere and on precipitation rates to model the frequency of lightning strikes across the continental United States. They predict that the number of lightning strikes will increase by about 12% for every degree of rise in global average air temperature. — HJS

*Science*, this issue p. 851

### CIRCADIAN RHYTHM

#### The Sun rises (and sets) on hamster memories

Internal circadian clocks set a rhythm to which biological systems beat. In addition to regulating sleep and wake cycles, the circadian system influences

learning and memory. Using Siberian hamsters as a model, Fernandez *et al.* evaluated the role of the part of the brain that regulates the circadian cycle, the suprachiasmatic nucleus (SCN), in altered memory processing when circadian rhythms are disrupted. As expected, animals with disrupted rhythms had recognition and spatial memory impairments. However, when the animals' SCN was disconnected, these memory deficits were reversed. Thus, memory impairment resulting from circadian dysfunction is dependent on preserving SCN circuit connections. — MM

*Science*, this issue p. 854



## EXOPLANET ATMOSPHERE

### Sunny side hot for tidally locked world

Most planets in our solar system spread their heat evenly across their surfaces in the course of a day. They accomplish this with relatively fast rotations and a generous distance from the Sun. Some exoplanets, however, aren't so well balanced. Stevenson *et al.* show that planets like the Jupiter-sized WASP-43b whirl closely around their star in less than 24 hours, which leaves them tidally locked with little chance for heat redistribution. — MMM

*Science*, this issue p. 838

## CILIA AND FLAGELLA

### Molecular ruler rules cilia and flagella length

Cilia and flagella contain within their ultrastructure repeating structures at regularly spaced intervals. How does the cell measure length with nanometer precision? Oda *et al.* identify a flagella protein complex in *Chlamydomonas* that appears to act as a sort of molecular

ruler to define repeat length. Genetic changes that would change the length of this protein led to corresponding changes in the length of repeats within the resulting flagella. — SMH

*Science*, this issue p. 857

## NANOPHOTONICS

### Gold nanoparticles form potential plasmons

Plasmons—the collective light-induced electronic excitations that occur at the surface of a metal—may form the basis for a new technology for harvesting solar energy. Currently, solar energy is converted into useful energy either by solar-thermal and photovoltaic conversion systems seen on rooftops, or, in some niche applications, by thermoelectric devices. Sheldon *et al.* observed electric potentials induced in gold nanospheres by optical irradiation. The effect may be useful in the design of a new generation of energy conversion devices based on plasmonics. — ISO

*Science*, this issue p. 828

## IMMUNOLOGY

### A step toward better vaccine adjuvants

The receptor TLR4 stimulates immune signaling pathways. It can do so through two adaptor proteins: MyD88, which can trigger undesirable inflammatory responses; and TRIF, which stimulates immune responses. Currently, adjuvants to boost immune responses to vaccines are developed with the idea that their structure determines the adaptor protein that TLR4 will use. However, Kolb *et al.* suggest that TLR4 signaling is inherently biased toward the TRIF-dependent pathway, particularly in the context of type I interferon signaling. The findings may help in the development of more effective vaccine adjuvants that enhance immune responses without triggering potentially harmful inflammatory reactions. — JFF

*Sci. Signal.* **7**, ra108 (2014).



PHOTOS: (LEFT TO RIGHT) ANDREW SYRED/SCIENCE SOURCE; © BLEND IMAGES/ALAMY

## IN OTHER JOURNALS

Edited by **Kristen Mueller**  
and **Jesse Smith**



## CIRCADIAN RHYTHMS

### One clock for you and your microbes

**D**isrupting our circadian rhythms increases the risk of developing diabetes, obesity, cancer, and cardiovascular disease, but scientists do not fully understand why. Thaiss *et al.* now report that conditions that cause jet lag change the composition and activities of gut microbes in mice, which can lead to metabolic disease. Gut microbe composition no longer fluctuated diurnally in mice with disrupted circadian rhythms, but normal rhythmic feeding or the transplantation of gut microbes from normal mice restored this oscillation. Normal mice that received gut microbial transplants from jet-lagged humans or mice that experienced a change in their day-night schedule gained weight and developed symptoms of metabolic disease. — LC

*Cell* **159**, 514 (2014).

## BIOTECHNOLOGY

### Aiming for a better carbon fix in plants

Enhancing the output of Rubisco, an enzyme that converts atmospheric CO<sub>2</sub> into energy-rich molecules, could improve photosynthetic efficiency, and therefore crop yield, in plants. Maize is a C4 grass, which uses four-carbon compounds to carry CO<sub>2</sub> into an interior compartment; subsequent release of CO<sub>2</sub> increases its local concentration and favors efficient activity of Rubisco.

Rice, however, is a C3 grass and lacks this pathway. Wang *et al.* compared transcripts and metabolites in developing maize and rice plants as a step toward understanding the biochemical and anatomical bases of C4 photosynthesis. Furthermore, Lin *et al.* transplanted Rubisco from a cyanobacterium, which also relies on a CO<sub>2</sub>-concentrating apparatus, into tobacco (a C3 plant) chloroplasts. — GJC

*Nat. Biotech.* **32**, 10.1038/nbt.3019 (2014); *Nature* **513**, 547 (2014).



## ALSO IN SCIENCE JOURNALS

Edited by Stella Hurtley

## PALEOCEANOGRAPHY

## Leading a wintry march from a distance

When the Antarctic ice sheet began to expand around 3 million years ago, it caused changes in deep ocean circulation, hastening the pace of glaciation in the Northern Hemisphere. Woodard *et al.* analyzed marine sediments from the northwest Pacific Ocean. A substantial fraction of the sea level fall actually occurred before the Northern Hemispheric ice sheets began to grow rapidly, probably because of continental ice growth in Antarctica. Thus, Antarctic glaciers appear to be more dynamic than anticipated, which has implications for the stability of the Antarctic ice sheet in a warming world. — HJS

*Science*, this issue p. 847

## SYNTHETIC BIOLOGY

## Record your memories with a DNA recorder

DNA-based memory devices are not optimal for recording analog information, such as the magnitude of inputs over time. Farzadfard and Lu converted genomic DNA into a "tape recorder" within living bacterial populations (see the Perspective by Ausländer and Fussenegger). Specific DNAs were used to introduce precise mutations into genomic DNA. The stored information could be read out via reporter genes, functional assays, and DNA sequencing. This approach allowed the memorization of multiple inputs at the population level. The record could also be erased when required. — BJ

*Science*, this issue p. 825; see also p. 813

## MESENCHYMAL CELLS

## How lung mesenchymal cells behave

Despite the variety of organ systems, there is a common theme: Stromal tissues support and maintain most vertebrate organs. These stromal tissues form from mesenchymal stem cells. Kumar *et al.* used clonal cell labeling in mice to identify and characterize stromal progenitors in the developing mouse lung at single-cell resolution (see the Perspective by Lee and Kim). Progenitor populations occupied different locations and displayed a variety of movements and lineage boundaries. Airway smooth muscle progenitors are located just ahead of budding branches in the bronchial tree and are organized into carefully controlled domains. — BAP

*Science*, this issue p. 827; see also p. 810

## DNA RECOMBINATION

## Mapping recombination in individual human males

Sperm and eggs form from diploid cells that contain two copies of our genomic DNA. The haploid germ cells must undergo a special cell division, meiosis, which halves their DNA content. Meiosis involves a DNA recombination step between parental chromosomes. Recombination is initiated by a DNA double-strand break, which can exchange DNA between the chromosomes, a process that drives human genetic variation. Pratto *et al.* mapped meiotic recombination sites in individual human males (see the Perspective by de Massy). Recombination hotspots were influenced by variants of the histone-lysine N-methyltransferase protein,

PRDM9, as well as by other factors. The recombination sites also influence genome evolution and the incidence of genetic disease. — GR

*Science*, this issue p. 826; see also p. 808

## INTERFACIAL WATER

## Dissecting the electrical double layer

The structure of water within a nanometer of an electrode surface is known as the electrical double layer. This layer creates a strong electrical field that can affect electrochemical reactions. Velasco-Velez *et al.* explored the structure of the electrical double layer at a bare gold electrode. With no applied potential and at positive potentials, the layer is highly structured (resembling ice) with few dangling hydrogen bonds. However, at negative potentials, the layer was more like bulk water, and half of the water molecules lie flat on the surface. — PDS

*Science*, this issue p. 831

## STRONG BASES

## Bring your own template for deprotonation

When manufacturing pharmaceuticals and agrochemicals, chemists need to add substituents to specific carbon sites in hexagonal benzene rings. If there's already a substituent on the ring, it can often direct a base to deprotonate the site next to it. But what if you want the base to attack the site two carbons away? Martínez-Martínez *et al.* devised a method to do this by taking advantage of the sodium and magnesium counterions associated with their base. These ions form a template that orients

the base to attack the more distant site. — JSY

*Science*, this issue p. 834

## LANDSCAPE HYDROLOGY

## Impacts of deforestation on wetlands

Deforestation worldwide may be causing an increase in the extent of wetlands. Using a combination of different approaches, Woodward *et al.* show that ancient and more recent deforestation has resulted in major changes in global wetland hydrology. For example, in Australia and New Zealand, deforestation has created new wetlands or increased the water level in existing wetlands. Recognition of this effect has implications for landscape management: Reforestation programs in wetland catchments may have unintended consequences for vulnerable wetlands. — AMS

*Science*, this issue p. 844

## ANTIBIOTIC RESISTANCE

## It takes two antibiotics—but which two?

Treatment with multiple antibiotics reduces the development of antibiotic-resistant bacteria, but not all combinations are created equal. Predictive models are needed to design therapeutic regimens that avoid resistance. Munck *et al.* used adaptive evolution and genomics to unravel resistance responses of bacteria to single and pairs of antibiotics. The authors pinpointed mutations that lead to different levels of drug resistance. The findings can be exploited to design better drug combinations that limit the acquisition of antibiotic resistance. — KL

*Sci. Transl. Med.* **6**, 262ra156 (2014).



Experiment site in a mixed-grass prairie near Cheyenne, Wyoming.

#### CLIMATE CHANGE ECOLOGY

### How elevated CO<sub>2</sub> affects grassland

**E**ven after two decades, experiments to study the ecological effects of climate change continue to yield instructive results. Zelikova *et al.* studied the effects of elevated CO<sub>2</sub> on grassland in Wyoming over 8 years and found that the plant productivity stabilized over time. The abundance of two dominant plant species (western wheat grass and blue grama grass) decreased with elevated CO<sub>2</sub>, which led to the increased presence of subdominant species and hence greater evenness in community composition. Plant productivity stabilized as a result of such alterations community structure. This research highlights the potential for plant communities to change in the face of changing climates. — AMS

*Proc. Natl. Acad. Sci. U.S.A.* **111**, 15456 (2014).

#### SUPERCONDUCTIVITY

### Describing iron-based superconductors

In a superconducting material, pairs of electrons (the so-called Cooper pairs) flow effortlessly through the material without encountering any resistance. The energy needed to break up a Cooper pair is called the superconducting gap. Since their discovery several years ago, iron-based superconductors (IBSs) have puzzled researchers because different IBS families appear to have different gap symmetries. Yin *et al.* used first-principles calculations to explore the nature of superconductivity in a large number of IBS compounds. They found that three related

types of symmetry occurred in different IBS families, including a variant that hadn't been discussed previously. A comparison with experimental data for the compound LiFeAs suggests that the gap in this material has this particular symmetry. — JS

*Nat. Phys.* **10**, 845 (2014).

#### OPEN INNOVATION

### Sharing results midway or at the end

Technological innovation depends on striking a balance between sharing discoveries to spur follow-on development and ensuring that too much sharing does not undercut incentives to participate. Boudreau and Lakhani randomized an online

bioinformatics algorithm contest in order to examine the consequences of intermediate and final disclosure. Final-disclosure (FD) solvers had no in-group communication, whereas intermediate-disclosure (ID) solvers could share intermediate solutions. Fewer people opted to participate in ID, they submitted fewer solutions per person, and they tried fewer unique approaches. However, ID approaches clustered together conceptually and smoothly improved more quickly over time, whereas FD had much more variable quality and didn't achieve the mean or peak performance levels shown in ID. — BW

*Res. Policy* 10.1016/j.respol.2014.08.001 (2014).

#### BIOFUELS

### A team effort to get more out of lignin

You've probably heard about cellulosic biofuels—fuels manufactured from the sugars trapped in plants' cellulose, which is largely inedible (for humans anyway). As commercial efforts in this arena get off the ground, chemists are rushing to solve a related problem: how to transform and market the woody lignin material that comes along with the cellulose. Linger *et al.* demonstrate the advantages of combining chemical and microbial protocols. First, treatment with base breaks the lignin down into a diverse set of molecules. Then *Pseudomonas putida* bacteria are put to work channeling these molecules into a more uniform product stream of (poly)-hydroxyalkanoates. After that, synthetic chemistry transforms this stream into plastics, liquid fuels, and commodity compounds. — JSY

*Proc. Natl. Acad. Sci. U.S.A.* 10.1073/pnas.1410657111 (2014).

#### TISSUE REPAIR

### A conserved response for tissue repair

Upon injury or infection, the body releases chemicals that trigger tissue repair and pathogen clearance. Because the medical community needs new therapeutic leads in this era of growing antibiotic resistance, identifying these molecules is a high priority. Dalli *et al.* looked for these factors in mice infected with self-resolving *Escherichia coli*, in human breast milk, and in regenerating planaria. They identified two related molecules, conserved across these organisms, which promoted pathogen clearance, reduced inflammation, and accelerated tissue regeneration. Scientists will need to carry out further studies to determine whether these chemicals have similar properties in humans. — BAP

*Proc. Natl. Acad. Sci. U.S.A.* 10.1073/pnas.1415006111 (2014).

PHOTO: JULIE KRAY/USDA ARS IN FORT COLLINS



## RESEARCH ARTICLE SUMMARY

## SYNTHETIC BIOLOGY

# Genomically encoded analog memory with precise in vivo DNA writing in living cell populations

Fahim Farzadfard and Timothy K. Lu\*

**INTRODUCTION:** The conversion of transient information into long-lasting responses is a common aspect of many biological processes and is crucial for the design of sophisticated synthetic circuits. Genomic DNA provides a rich medium for the storage of information in living cells. However, current cellular memory technologies are limited in their storage capacity and scalability.

**RATIONALE:** We converted genomic DNA into a “tape recorder” for memorizing information in living cell populations. This was achieved via SCRIBE (Synthetic Cellular Recorders Integrating Biological Events), a programmable and modular architecture for generating single-stranded DNA (ssDNA) inside of living cells in response to gene regulatory signals. When coexpressed with a recombinase, these ssDNAs address specific target loci on the basis of sequence homology and introduce precise mutations into genomic DNA, thus converting transient cellular signals into genomically encoded memory. This distributed biological memory leverages the large number of cells in bacterial cultures and encodes information into their collective genomic DNA in the form of the fraction of cells that carry specific mutations.

**RESULTS:** We show that SCRIBE enables the recording of arbitrary transcriptional inputs into DNA storage registers in living cells by translating regulatory signals into ssDNAs. In *Escherichia coli*, we expressed ssDNAs from engineered retransons that use a reverse transcriptase protein to produce hybrid RNA-ssDNA molecules. These intracellularly expressed ssDNAs are targeted

into specific genomic loci where they are recombined and converted into permanent memory. We show that genomically stored information can be readily reprogrammed by changing the ssDNA template and controlled via both chemical and light inputs. We demonstrate that genomically encoded memory can be read with a variety of techniques, including reporter genes, functional assays, and high-throughput DNA sequencing.

SCRIBE enables the recording of analog information such as the magnitude and time span of exposure to an input. This convenient feature is facilitated by the intermediate recombination rate of our current system ( $\sim 10^{-4}$  recombination events per generation), which we validated via a mathematical model and computer simulations. For example, we stored the overall exposure time to chemical inducers in the

DNA memory of bacterial populations for 12 days ( $\sim 120$  generations), independently of the induction pattern. The frequency of mutants in these populations was linearly related to the total exposure time.

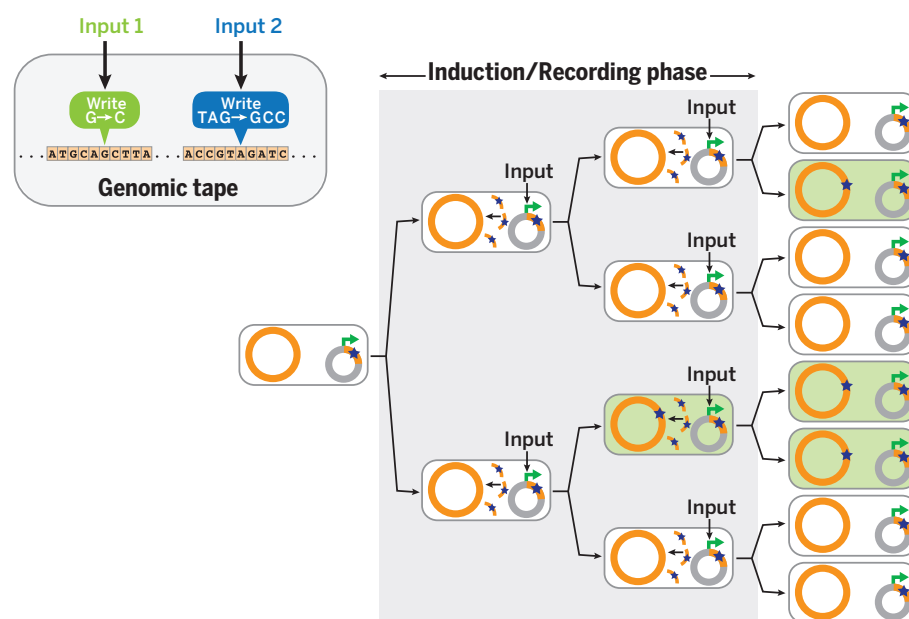
Furthermore, we demonstrate that SCRIBE-induced mutations can be written and erased and can be used to record multiple inputs across the distributed genomic DNA of bacterial populations. Finally, we show that SCRIBE memory can be decomposed into independent “input,” “write,” and “read” operations and used to create genetic “logic-and-memory” circuits, as well as “sample-and-hold” circuits.

**CONCLUSION:** We describe a scalable platform that uses genomic DNA for analog, rewritable, and flexible memory distributed across living cell populations. We anticipate that SCRIBE

## ON OUR WEB SITE

Read the full article at <http://dx.doi.org/10.1126/science.1256272>

will enable long-term cellular recorders for environmental and biomedical applications. Future optimization of recombination efficiencies achievable by SCRIBE could lead to more efficient single-cell digital memories and enhanced genome engineering technologies. Furthermore, the ability to regulate the generation of arbitrary targeted mutations with other gene-editing technologies should enable genomically encoded memory in additional organisms. ■



**SCRIBE enables distributed genomically encoded memory.** In the presence of an input, ssDNAs (orange curved lines) are produced from a plasmid-borne cassette (gray circles) and recombined into specific genomic loci (orange circles) that are targeted on the basis of sequence homology. This results in the accumulation of precise mutations (stars in green cells) as a function of the magnitude and duration of exposure to the input.

## RELATED ITEMS IN SCIENCE

S. Ausländer, M. Fussenegger, Dynamic genome engineering in living cells. *Science* **346**, 813–814 (2014).

The list of author affiliations is available in the full article online.

\*Corresponding author. E-mail: [timlu@mit.edu](mailto:timlu@mit.edu)

Cite this article as F. Farzadfard, T. K. Lu, *Science* **346**, 1256272 (2014). DOI: 10.1126/science.1256272

## RESEARCH ARTICLE

## SYNTHETIC BIOLOGY

# Genomically encoded analog memory with precise in vivo DNA writing in living cell populations

Fahim Farzadfard<sup>1,2,3</sup> and Timothy K. Lu<sup>1,2,3\*</sup>

Cellular memory is crucial to many natural biological processes and sophisticated synthetic biology applications. Existing cellular memories rely on epigenetic switches or recombinases, which are limited in scalability and recording capacity. In this work, we use the DNA of living cell populations as genomic “tape recorders” for the analog and distributed recording of long-term event histories. We describe a platform for generating single-stranded DNA (ssDNA) in vivo in response to arbitrary transcriptional signals. When coexpressed with a recombinase, these intracellularly expressed ssDNAs target specific genomic DNA addresses, resulting in precise mutations that accumulate in cell populations as a function of the magnitude and duration of the inputs. This platform could enable long-term cellular recorders for environmental and biomedical applications, biological state machines, and enhanced genome engineering strategies.

Due to its high storage capacity, durability, ease of duplication, and high-fidelity maintenance of information, DNA has garnered much interest as an artificial storage medium (1, 2). However, existing technologies for in vivo autonomous recording of information in cellular memory are limited in their storage capacity and scalability (3). Epigenetic memory devices such as bistable toggle switches (4–7) and positive-feedback loops (8) require orthogonal transcription factors and can lose their digital state due to environmental fluctuations or cell death. Recombinase-based memory devices enable the writing and storage of digital information in the DNA of living cells (9–12), where binary bits of information are stored in the orientation of large stretches of DNA. However, these devices do not efficiently exploit the full capacity of DNA for information storage: Recording a single bit of information with these devices often requires at least a few hundred base pairs of DNA, overexpression of a recombinase protein to invert the target DNA, and engineering recombinase recognition sites into target loci in advance. The scalability of this type of memory is further limited by the number of orthogonal recombinases that can be used in a single cell. Finally, epigenetic and recombinase-based memory devices described to date store digital information, and their recording capacity is exhausted within a

few hours of induction. Thus, these devices have not been adapted to record analog information, such as magnitude and time course of inputs over extended periods of time (i.e., multiple days or more).

Here we introduce SCRIBE (Synthetic Cellular Recorders Integrating Biological Events), a compact, modular strategy for producing single-stranded DNA (ssDNA) inside of living cells in response to a range of regulatory signals, such as small chemical inducers and light. These ssDNAs address specific target loci on the basis of sequence homology and introduce precise mutations into genomic DNA. The memory device can be easily reprogrammed to target different genomic locations by changing the ssDNA template. SCRIBE memory does not just record the absence or presence of arbitrary inputs (digital signals represented as binary “0s” or “1s”). Instead, by encoding information into the collective genomic DNA of cell populations, SCRIBE can track the magnitude and long-term temporal behavior of inputs, which are analog signals that can vary over a wide range of continuous values. This analog memory architecture leverages the large number of cells in bacterial cultures for distributed information storage and archives event histories in the fraction of cells in a population that carry specific mutations.

## Single-stranded DNA expression in living cells

Previously, it was shown that synthetic oligonucleotides delivered by electroporation into cells that overexpress Beta recombinase (from bacteriophage  $\lambda$ ) in *Escherichia coli* are specifically and efficiently recombined into homologous genomic sites (13–16). Thus, oligonucleotide-mediated recombination offers a powerful way to intro-

duce targeted mutations in a bacterial genome (17, 18). However, this technique requires the exogenous delivery of ssDNAs and cannot be used to couple arbitrary signals into genetic memory. To overcome these limitations, we developed a genome editing platform based on expressing ssDNAs inside of living cells by taking advantage of a widespread class of bacterial reverse transcriptases (RTs) called retrons (19, 20).

The wild-type (WT) retron cassette encodes three components in a single transcript: a RT protein and two RNA moieties, *msr* and *msd*, which act as the primer and the template for the reverse transcriptase, respectively (Fig. 1A, left). The *msr-msd* sequence in the retron cassette is flanked by two inverted repeats. Once transcribed, the *msr-msd* RNA folds into a secondary structure guided by the base pairing of the inverted repeats and the *msr-msd* sequence. The RT recognizes this secondary structure and uses a conserved guanosine residue in the *msr* as a priming site to reverse transcribe the *msd* sequence and produce a hybrid RNA-ssDNA molecule called msDNA (i.e., multicopy single-stranded DNA) (20, 21). To couple the expression of ssDNA to an external input, the WT Ec86 retron cassette from *E. coli* BL21 (21) was placed under the control of an isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG)-inducible promoter ( $P_{lacO}$ ) in *E. coli* DH5aPRO cells (22), which express high levels of the LacI and TetR repressors (Fig. 1A). The WT retron ssDNA [ssDNA(wt)] was readily detected in IPTG-induced cells, whereas no ssDNA was detected in noninduced cells (Fig. 1B). The identity of the detected ssDNA band was further confirmed by DNA sequencing (fig. S1). To verify that ssDNA expression depended on RT activity, point mutations [Asp<sup>197</sup>→Ala<sup>197</sup> (D197A) and D198A] were introduced to the active site of the RT to make a catalytically dead RT (dRT) (23). This modification completely abolished ssDNA production (Fig. 1B).

To engineer the *msd* template to express synthetic ssDNAs of interest, we initially tried to replace the whole *msd* sequence with a desired template. However, no ssDNA was detected, suggesting that some features of *msd* are required for ssDNA expression, as was previously noted for another retron (24). A variant in which the flanking regions of the *msd* stem remained intact (Fig. 1A, right) produced detectable amounts of ssDNA when induced by IPTG (Fig. 1B,  $P_{lacO\_msd}(kanR)_{ON}$  + IPTG). The correct identity of the detected ssDNA band was further confirmed by DNA sequencing (fig. S1). Thus, the lower part of the *msd* stem is essential for reverse transcription, whereas the upper part of the stem and the loop are dispensable and can be replaced with desired templates to produce ssDNAs of interest in vivo.

## Regulated genome editing with in vivo ssDNAs

To demonstrate that intracellularly expressed ssDNAs can be recombined into target genomic loci by concomitant expression of Beta, we developed a selectable marker reversion assay (Fig. 1C). The *kanR* gene, which encodes neomycin

<sup>1</sup>Synthetic Biology Group, Research Laboratory of Electronics, Department of Electrical Engineering and Computer Science and Department of Biological Engineering, Massachusetts Institute of Technology (MIT), 77 Massachusetts Avenue, Cambridge, MA 02139, USA. <sup>2</sup>MIT Synthetic Biology Center, 500 Technology Square, Cambridge, MA 02139, USA. <sup>3</sup>MIT Microbiology Program, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. \*Corresponding author. E-mail: timlu@mit.edu



phosphotransferase II and confers resistance to kanamycin (Kan), was integrated into the *galK* locus. Two stop codons were then introduced into the genomic *kanR* to make a Kan-sensitive *kanR<sub>OFF</sub>* reporter strain (DH5aPRO *galK::kanR<sub>W28TAA, A29TAG</sub>*). These premature stop codons could be reverted back to the WT sequence via recombination with engineered ssDNA(*kanR*)<sub>ON</sub>, thus conferring kanamycin resistance (Fig. 1C). Specifically, ssDNA(*kanR*)<sub>ON</sub> contains 74 base pairs of homology to the regions of the *kanR<sub>OFF</sub>* locus flanking the premature stop codons and replaces the stop codons with the WT *kanR* gene sequence (Fig. 1C).

We cloned the Beta gene (*bet*) into a plasmid under the control of the anhydrotetracycline (aTc)-inducible *P<sub>tetO</sub>* promoter and introduced it along with the IPTG-inducible *msd(kanR)*<sub>ON</sub> construct into the *kanR<sub>OFF</sub>* strain (Fig. 1C). Induction of cultures harboring these two plasmids with either IPTG (1 mM) or aTc (100 ng/ml) resulted in a slight increase in the frequency of Kan-resistant cells within the population (Fig. 1C). However, coexpression of both ssDNA(*kanR*)<sub>ON</sub> and Beta with IPTG and aTc resulted in a  $>10^4$ -fold increase in the recombinant frequency relative to the non-induced cells. This corresponded to a  $>10^3$ -fold

increase relative to cells induced with IPTG only and a 60-fold increase relative to cells induced with aTc only. This increase in the recombinant frequency was dependent on the RT activity, as it was largely abolished with dRT. The genotypes of randomly selected Kan-resistant colonies were further confirmed by DNA sequencing to contain precise reversions of the two codons to the WT sequence (fig. S1). No Kan-resistant colonies were detected when a nonspecific ssDNA [ssDNA(wt)] was coexpressed with Beta in the *kanR<sub>OFF</sub>* reporter cells, confirming that Kan-resistant cells were not produced due to spontaneous mutations. In additional experiments, we used high-throughput sequencing (Illumina HiSeq) on the bacterial populations to analyze the genomically encoded memory (see supplementary materials and fig. S2). Comparable recombinant frequencies were obtained from both the plating assay and sequencing, confirming that genomically encoded memory can be read without the need for functional assays and reporters.

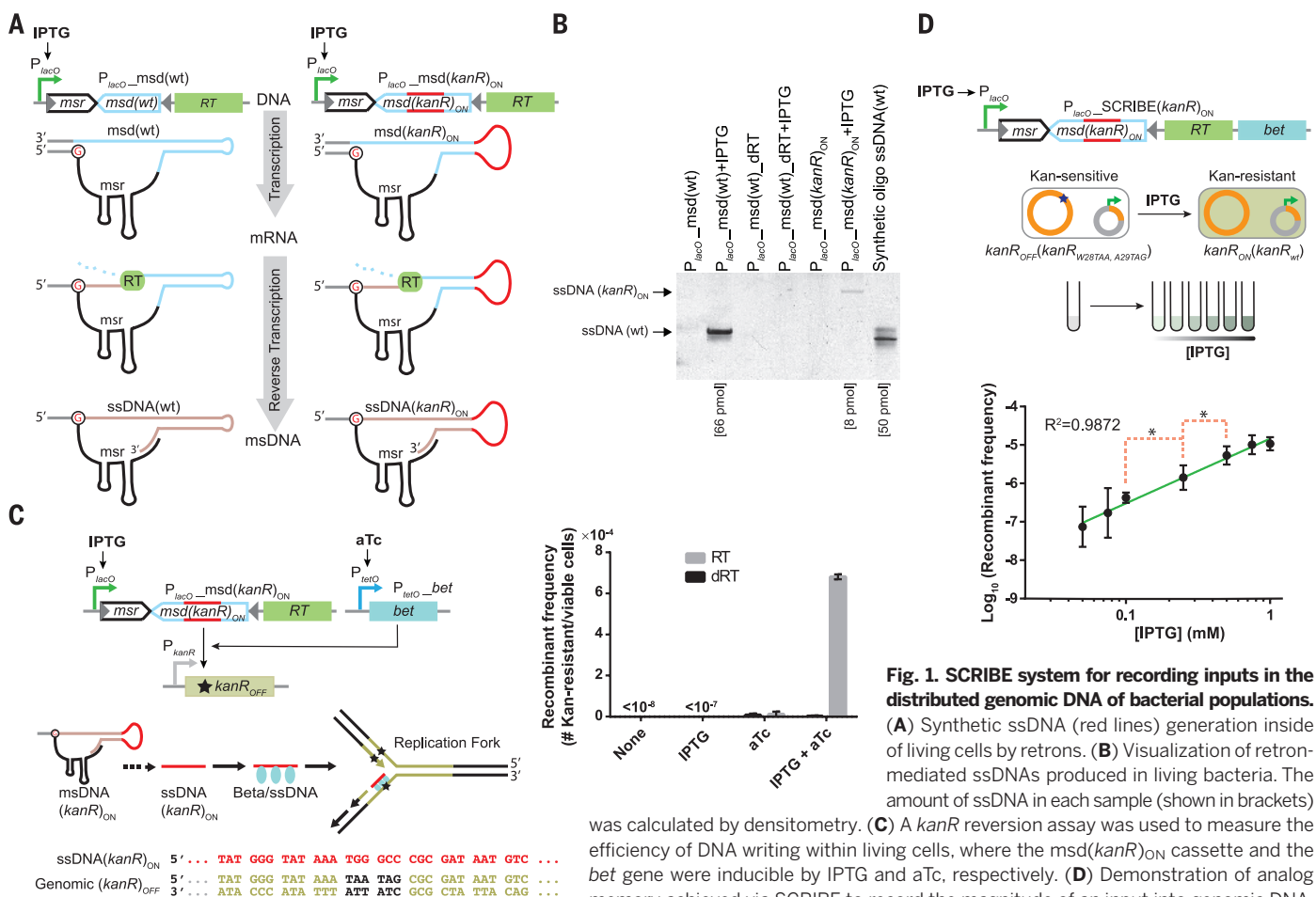
### Recording input magnitudes into genomic memory

We reasoned that the rate of recombination between engineered ssDNAs and genomic DNA

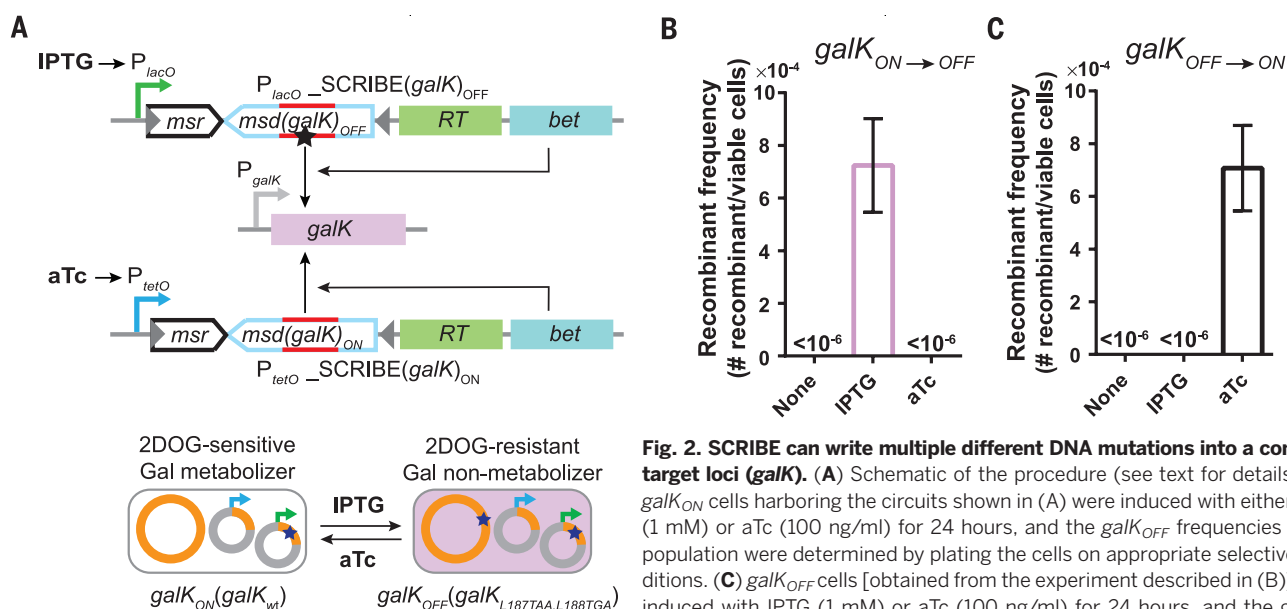
could be effectively modulated by changing expression levels of the engineered retron cassette and Beta. This feature would enable the recording of analog information, such as the magnitude of an input signal, in the proportion of cells in a population with a specific mutation in genomic DNA. To demonstrate this, both the ssDNA(*kanR*)<sub>ON</sub> expression cassette and *bet* were placed into a single synthetic operon [hereafter referred to as the SCRIBE(*kanR*)<sub>ON</sub> cassette] under the control of *P<sub>lacO</sub>* (Fig. 1D). The *kanR<sub>OFF</sub>* reporter cells harboring this synthetic operon were induced with different concentrations of IPTG. The fraction of Kan-resistant recombinants increased linearly with the input inducer concentration on a log-log plot over a range of  $\sim 10^{-7}$  to  $\sim 10^{-5}$  (Fig. 1D). Statistical tests showed that at least four different concentrations of the inducer (including 0 mM IPTG) could be resolved in this experiment. Thus, the efficiency of genome writing in a population can be quantitatively tuned with external inputs.

### Writing and rewriting genomic memory

We next created a complementary set of SCRIBE cassettes to write and erase (rewrite) information in the genomic *galK* locus using two different chemical inducers. Cells expressing *galK* can



The green line is a linear regression fit. The red dashed brackets marked with asterisks connect the closest data points that are statistically significant with respect to each other ( $P < 0.05$  based on one-tailed Welch's *t* test). Error bars indicate SEM for three independent biological replicates.



**Fig. 2. SCRIPE can write multiple different DNA mutations into a common target loci (*galk*).** (A) Schematic of the procedure (see text for details). (B)  $galk_{ON}$  cells harboring the circuits shown in (A) were induced with either IPTG (1 mM) or aTc (100 ng/ml) for 24 hours, and the  $galk_{OFF}$  frequencies in the population were determined by plating the cells on appropriate selective conditions. (C)  $galk_{OFF}$  cells [obtained from the experiment described in (B)] were induced with IPTG (1 mM) or aTc (100 ng/ml) for 24 hours, and the  $galk_{ON}$  frequencies in the population were determined by plating the cells on appropriate selective conditions. Error bars indicate SEM for three independent biological replicates.

metabolize and grow on galactose as the sole carbon source. However, these  $galk$ -positive ( $galk_{ON}$ ) cells cannot metabolize 2-deoxy-galactose (2DOG) and cannot grow on plates containing glycerol (carbon source) + 2DOG. On the other hand,  $galk$ -negative ( $galk_{OFF}$ ) cells cannot grow on galactose as the sole carbon source but can grow on glycerol + 2DOG plates (25). We transformed DH5 $\alpha$ PRO  $galk_{ON}$  cells with plasmids encoding IPTG-inducible SCRIPE( $galk$ ) $_{OFF}$  and aTc-inducible SCRIPE( $galk$ ) $_{ON}$  cassettes (Fig. 2A). Induction of SCRIPE( $galk$ ) $_{OFF}$  by IPTG resulted in the writing of two stop codons into  $galk_{ON}$ , leading to  $galk_{OFF}$  cells that could grow on glycerol + 2DOG plates (Fig. 2B). Induction of SCRIPE( $galk$ ) $_{ON}$  in these  $galk_{OFF}$  cells with aTc reversed the IPTG-induced modification, leading to  $galk_{ON}$  cells that could grow on galactose plates (Fig. 2C). These results show that writing on genomic DNA with SCRIPE is reversible and that distinct information can be written and rewritten into the same locus.

### Writing multiple mutations into independent loci

Scaling the capacity of previous memory devices is challenging because each additional bit of information requires additional orthogonal proteins (e.g., recombinases or transcription factors). In contrast, orthogonal SCRIPE memory devices are potentially easier to scale because they can be built by simply changing the ssDNA template (*msd*). To demonstrate this, we used SCRIPE to record multiple independent inputs into different genomic loci of bacterial population. We integrated the  $kanR_{OFF}$  reporter gene into the *bioA* locus of DH5 $\alpha$ PRO to create a  $kanR_{OFF}galk_{ON}$  strain. These cells were then transformed with plasmids encoding IPTG-inducible SCRIPE( $kanR$ ) $_{ON}$  and aTc-inducible SCRIPE( $galk$ ) $_{OFF}$  cassettes (Fig. 3A). Induction of these cells with IPTG or aTc resulted

in the production of cells with phenotypes corresponding to  $kanR_{ON}galk_{ON}$  or  $kanR_{OFF}galk_{OFF}$  genotypes, respectively (Fig. 3, B and C). Comparable numbers of  $kanR_{ON}galk_{ON}$  and  $kanR_{OFF}galk_{OFF}$  cells ( $\sim 2 \times 10^{-4}$  and  $\sim 3 \times 10^{-4}$  recombinant/viable cells, respectively) were produced when the cultures were induced with both aTc and IPTG (Fig. 3C, left panel). Furthermore, very few individual colonies ( $\sim 3 \times 10^{-7}$  recombinant/viable cells) containing both writing events ( $kanR_{ON}galk_{OFF}$ ) were obtained in the cultures that were induced with both aTc and IPTG (Fig. 3C, right panel). These data suggest that although multiplexed writing at single-cell level is rare with SCRIPE's current level of recombination efficiency, multiple independent inputs can be successfully recorded into the distributed genomic DNA of bacterial subpopulations.

### Optogenetic genome editing for light-to-DNA memory

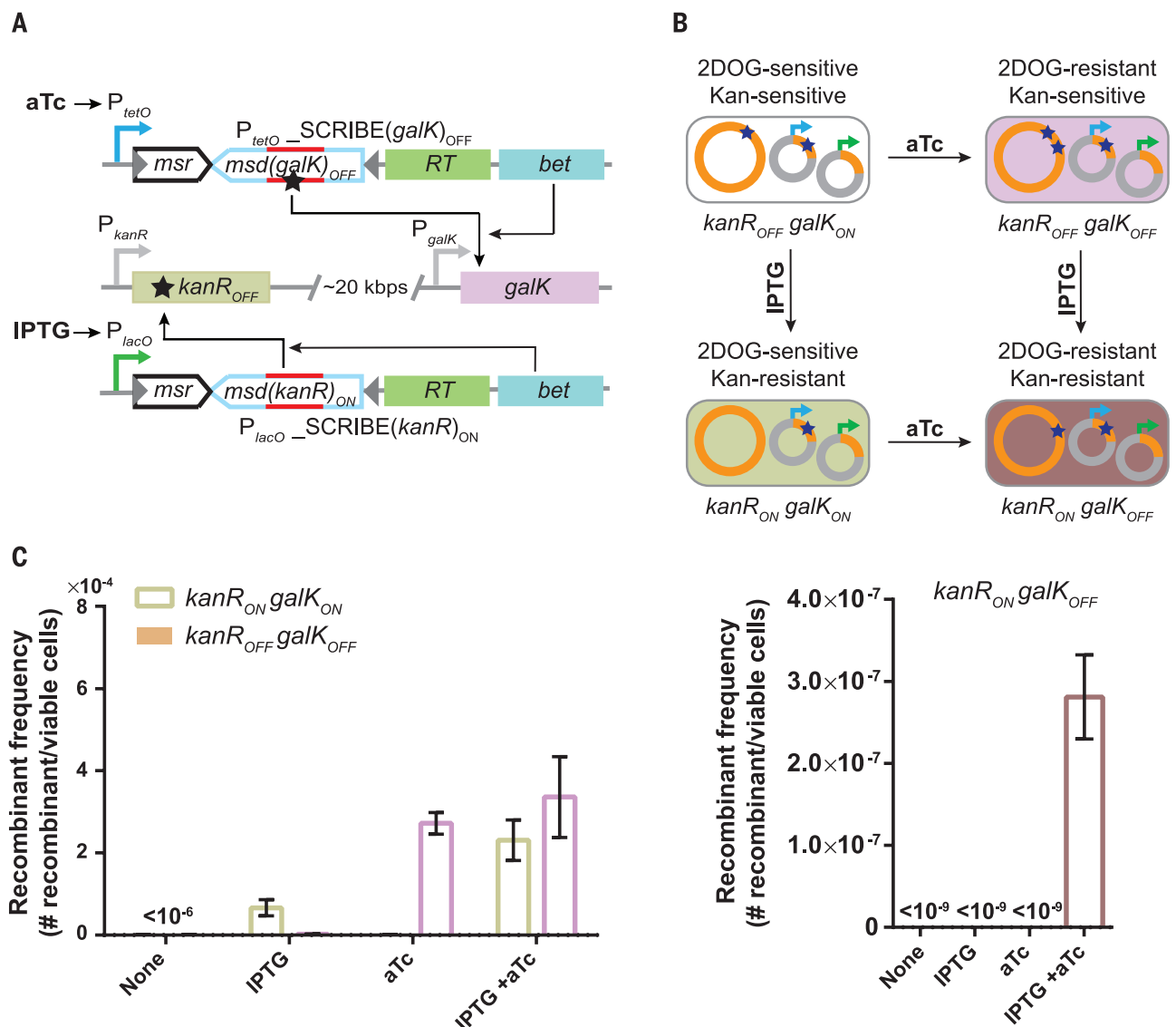
In SCRIPE, the expression of each individual ssDNA can be triggered by any endogenous or exogenous signal that can be coupled into transcriptional regulation, thus recording these inputs into long-lasting DNA storage. In addition to small-molecule chemicals, we showed that light can be used to trigger specific genome editing for genomically encoded memory. We placed the SCRIPE( $kanR$ ) $_{ON}$  cassette under the control of a previously described light-inducible promoter ( $P_{Dawn}$ ) (26) within  $kanR_{OFF}$  cells (Fig. 4A). These cultures were then grown for 4 days in the presence of light or in the dark (Fig. 4A). As Beta-mediated recombination is reportedly replication-dependent (27–29), dilutions of these cultures were made into fresh media at the end of each day to maintain active replication in the cultures. At the end of each day, samples were taken to determine the number of Kan-resistant and viable cells (Fig. 4A). Cultures grown in the dark yielded undetectable

levels of Kan-resistant cells (Fig. 4A). In contrast, the frequency of Kan-resistant cells increased steadily over time in the cultures that were grown in the presence of light, indicating the successful recording of light input into long-lasting DNA memory. The analog memory faithfully stored the total time of light exposure, rather than just the digital presence or absence of light.

### Recording the exposure time of inputs

The linear increase in the frequency of Kan-resistant colonies over time due to exposure to light indicates that the duration of inputs can be recorded into population-wide DNA memory using SCRIPE. To further explore population-wide genomically encoded memory whose state is a function of input exposure time, we used the  $kanR_{OFF}$  strain harboring the constructs shown in Fig. 1C, where expression of ssDNA( $kanR$ ) $_{ON}$  and Beta are controlled by IPTG and aTc, respectively. These cells were subjected to four different patterns of the inputs for 12 successive days (patterns I to IV, Fig. 4B). Kan-resistant cells did not accumulate in the negative control (pattern I), which was never exposed to the inducers. The fraction of Kan-resistant cells in the three other patterns (II, III, and IV) increased linearly over their respective induction periods and remained relatively constant when the inputs were removed. These data indicate that the genomically encoded memory was stable in the absence of the inputs over the course of the experiment. The recombinant frequencies in patterns III and IV, which were induced for the same total amount of time but with different temporal patterns, reached comparable levels at the end of the experiment. These data demonstrate that the genomic memory integrates over the total induction time and is independent of the input pattern, and therefore can be used to stably record event histories over many days.





**Fig. 3. Writing multiple mutations into independent target loci within population.** (A) Constructs used to target genomic *kanR*<sub>OFF</sub> and *galK*<sub>ON</sub> loci with IPTG-inducible and aTc-inducible SCRIBE cassettes, respectively. (B) Induction of *kanR*<sub>ON</sub> *galK*<sub>ON</sub> cells with IPTG or aTc generates cells with the *kanR*<sub>ON</sub> *galK*<sub>ON</sub> or *kanR*<sub>OFF</sub> *galK*<sub>OFF</sub> genotypes, respectively. Induction of *kanR*<sub>OFF</sub> *galK*<sub>ON</sub> cells with both IPTG

and aTc generates cells with the *kanR*<sub>ON</sub> *galK*<sub>OFF</sub> genotype. (C) *kanR*<sub>OFF</sub> *galK*<sub>ON</sub> reporter cells containing the circuits in (A) were induced with different combinations of IPTG (1 mM) and aTc (100 ng/ml) for 24 hours at 30°C, and the fraction of cells with the various genotypes were determined by plating the cells on appropriate selective media. Error bars indicate SEM for three independent biological replicates.

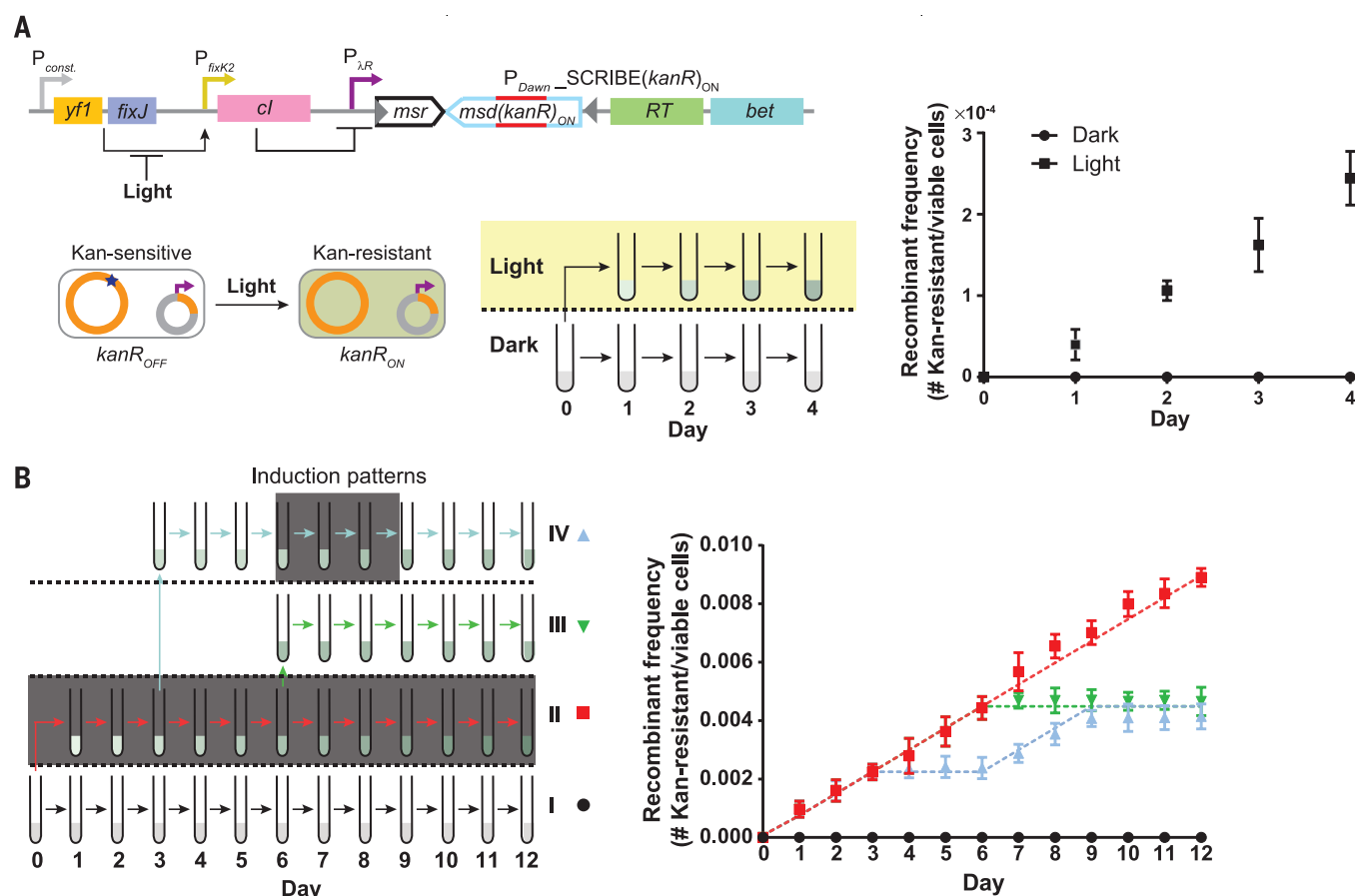
The linear increase in the fraction of recombinants in the induced cell populations over time was consistent with a deterministic model (dashed lines in Fig. 4B, also see supplementary materials). Specifically, when triggered by inputs, SCRIBE can substantially increase the rate of recombination events at a specific target site above the WT rate [which is reportedly <10<sup>-10</sup> events per generation in *recA*<sup>-</sup> background (30)]. When recombination rates are ~10<sup>-4</sup> events per generation, which is consistent with the recombination rate estimated for SCRIBE from data in Fig. 4B, a simple deterministic model and a detailed stochastic simulation both predict a linear increase in the frequency of recombinant alleles in a population over time, as long as this frequency is less than a few percent and cells in the population are equally fit over the time scale of interest (see

supplementary materials and figs. S3 and S4). These models enable one to determine the ideal range of recombination rates for a given application, which depends on parameters such as the frequency of dilution, the sensitivity of the method used for reading the memory, the desired input duration to be recorded, and so forth. For example, recombination rates that are too low would be challenging to quantify and could result in loss of memory if the cultures were diluted. Moreover, higher recombination rates lead to more rapid saturation of memory capacity in which the system is unable to provide a straightforward linear relation between the input exposure time and the state of the memory (fig. S3). Thus, intermediate levels of recombination rates are desirable for population-level analog memory units that can record the

time span of exposure to inputs (see supplementary materials).

### Decoupling memory operations

SCRIBE memory can be used to create more complex synthetic memory circuits. To demonstrate this, we first built a synthetic gene circuit that can record different input magnitudes into DNA memory. The memory state can then be read out later (after the initial input is removed) upon addition of a secondary signal. Specifically, we built an IPTG-inducible *lacZ*<sub>OFF</sub> (*lacZ*<sub>A35TAA, S36TAG</sub>) reporter construct in DH5αPRO cells (Fig. 5A). Expression of this reporter is normally repressed except when IPTG (“read” signal, Fig. 5A) is added, thus enabling a convenient and switchable population-level readout of the memory based on total LacZ activity (Fig. 5B). The *lacZ*<sub>OFF</sub>



**Fig. 4. Optogenetic genome editing and analog memory for long-term recording of input signal exposure times in the genomic DNA of living cell populations.** (A) We coupled expression of SCRIBE(*kanR*)<sub>ON</sub> to an optogenetic system ( $P_{Dawn}$ ). The *yf1/fixJ* synthetic operon was expressed from a constitutive promoter: In dark conditions, YF1 interacts with and phosphorylates FixJ. Phosphorylated FixJ activates the  $P_{fixK2}$  promoter, which drives  $\lambda$  repressor (*cl*) expression, which subsequently represses the SCRIBE(*kanR*)<sub>ON</sub> cassette. Light inhibits the interaction between YF1 and FixJ, leading to the generation of ssDNA(*kanR*)<sub>ON</sub> and Beta expression and, thus, the conversion of *kanR<sub>OFF</sub>* to *kanR<sub>ON</sub>*. Cells harboring this circuit were grown overnight at 37°C in the dark, diluted 1:1000, and then incubated for 24 hours at 30°C in the dark (no shading) or in the presence of light (yellow shading). Subsequently, cells were diluted by 1:1000

and grown for another 24 hours at 30°C in the dark or in the presence of light. The dilution-regrowth cycle was performed for four consecutive days. The *kanR* allele frequencies in the populations were determined by sampling the cultures after each 24-hour period. (B) SCRIBE analog memory records the total time exposure to a given input, regardless of the underlying induction pattern. Cells harboring the circuit shown in Fig. 1C were grown in four different patterns (I to IV) over a 12-day period, where induction by IPTG (1 mM) and aTc (100 ng/mL) is represented by dark gray shading. At the end of each 24-hour incubation period, cells were diluted by 1:1000 into fresh media. The frequency of Kan-resistant cells in the cultures was determined at the end of each day. Dashed lines represent the recombinant allele frequencies predicted by the model (see supplementary materials). Error bars indicate SEM for three independent biological replicates.

reporter cells were transformed with a plasmid encoding an aTc-inducible SCRIBE(*lacZ*)<sub>ON</sub> cassette (Fig. 5A). Overnight cultures were diluted and induced with various amounts of aTc to write the genomic memory (Fig. 5B). These cells were grown up to saturation and then diluted into fresh media in the presence or absence of IPTG to read the genomic memory (Fig. 5B). In the absence of IPTG, the total LacZ activity remained low, regardless of the aTc concentration. In the presence of IPTG, cultures that had been exposed to higher aTc concentrations had greater total LacZ activity. These results show that population-level reading of genomically encoded memory can be decoupled from writing and controlled externally. Furthermore, this circuit enables the magnitude of the inducer (aTc) to be stably recorded in the distributed genomic memory of a cellular population. Independent control over the read memory operation

as shown in this experiment could help to minimize fitness costs associated with the expression of reporter genes until needed.

We have shown that (i) both ssDNA expression and Beta are required for writing into genomic memory (Fig. 1C), (ii) multiple ssDNAs can be used to independently address different memory units (Fig. 3), and (iii) genomic memory is stably recorded into DNA and can be used to modify functional genes whose expression can be controlled by external inducers (Figs. 1 to 4). Thus, SCRIBE memory units can be conceptually decomposed into separate “input,” “write,” and “read” operations to facilitate greater control and the integration of logic with memory. The separation of these signals could enable master control over the writing of multiple independent inputs into genomic memory. To achieve this, we placed the *msd(lacZ)<sub>ON</sub>* cassette under the control of an acyl homoserine lactone (AHL)-inducible promoter

( $P_{luxR}$ ) (31) and cotransformed this plasmid with an aTc-inducible Beta-expressing plasmid into the *lacZ<sub>OFF</sub>* reporter strain (Fig. 5D). Using this design, information on the input (ssDNA expression via addition of AHL) can be written into DNA memory only in the presence of the write signal (Beta expression via addition of aTc). The information recorded in the memory register (i.e., the state of *lacZ* across the population) can be retrieved by adding the read signal (IPTG).

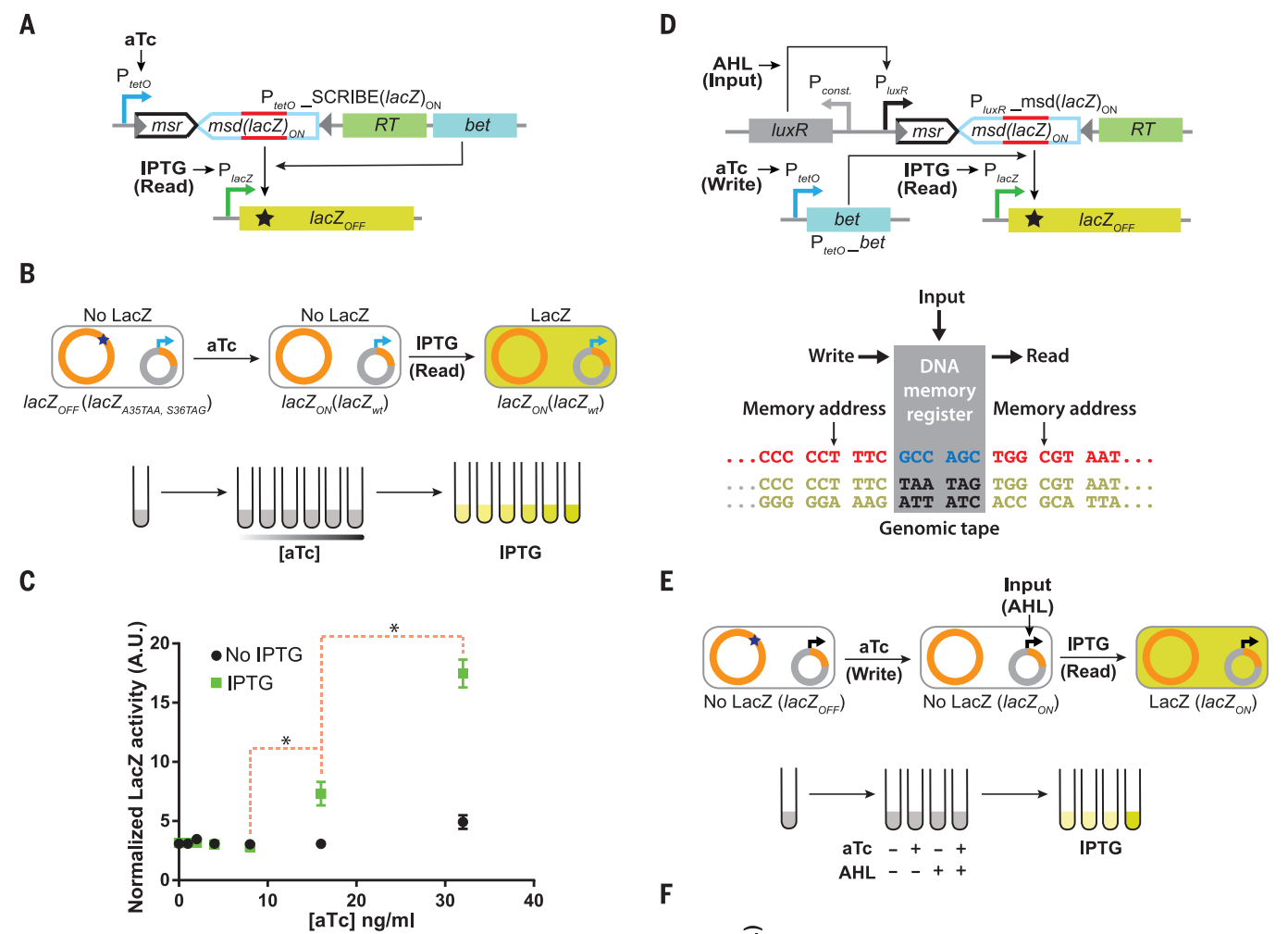
To demonstrate this, overnight *lacZ<sub>OFF</sub>* cultures harboring the circuit shown in Fig. 5D were diluted and then grown to saturation in the presence of all four possible combinations of AHL and aTc (Fig. 5E). The saturated cultures were then diluted into fresh media in the absence or presence of IPTG. As shown in Fig. 5F, only cultures that had been exposed to both the input and write signals simultaneously showed substantial LacZ activity, and only when they were



induced with the read signal. These results indicate that short stretches of DNA of living organisms can be used as addressable read/write memory registers to record transcriptional inputs.

Furthermore, SCRIBE memory can be combined with logic, such as the AND function between the input and write signals shown here. The logic in Fig. 5D enables this circuit to act as a “sample-

and-hold” system in which information about an input can be recorded in the presence of another signal and read out at will. Additional inputs in the form of orthogonal ssDNAs under the control



**Fig. 5. SCRIBE memory operations can be decoupled into independent “input,” “write,” and “read” operations, thus facilitating greater control over addressable memory registers in genomic tape recorders and the creation of sample-and-hold circuits.** (A) We built a circuit where information about the first inducer (aTc) is recorded in the population, which can then be read later upon addition of a second inducer (IPTG) that triggers a read operation. We created an IPTG-inducible *lacZ<sub>OFF</sub>* locus in the DH5αPRO background, which contains the full-length *lacZ* gene with two premature stop codons inside the open-reading frame. Expression of ssDNA(*lacZ*)<sub>ON</sub> from the aTc-inducible SCRIBE(*lacZ*)<sub>ON</sub> cassette results in the reversion of the stop codons inside *lacZ<sub>OFF</sub>* to yield the *lacZ<sub>ON</sub>* genotype. (B) Cells harboring the circuit shown in (A) were grown in the presence of different levels of aTc for 24 hours at 30°C to enable recording into genomic DNA. Subsequently, cell populations were diluted into fresh media without or with IPTG (1 mM) and incubated at 37°C for 8 hours. (C) Total LacZ activity in these cultures was measured using a fluorogenic *lacZ* substrate (FDG) assay. The red dashed brackets marked with asterisks connect the closest data points of IPTG-induced samples that are statistically significant ( $P < 0.05$  based on one-tailed Welch’s  $t$  test). A.U., arbitrary units. (D) We extended the circuit in (A) to create a sample-and-hold circuit where input, write, and read operations are independently controlled. This feature enables the creation of addressable read/write memory registers in the genomic DNA tape. Induction of cells with the input signal (AHL) produces ssDNA(*lacZ*)<sub>ON</sub>, which targets the genomic *lacZ<sub>OFF</sub>* locus for reversion to the WT sequence. In the presence of the write signal (aTc), which expresses Beta, ssDNA(*lacZ*)<sub>ON</sub> is recombined into the *lacZ<sub>OFF</sub>* locus and produces the *lacZ<sub>ON</sub>* genotype. Thus, the write signal enables the input signal to be sampled and held in memory. The total LacZ activity in the cell populations is retrieved by adding the read signal (IPTG). (E) Cells harboring the circuit shown in (D) were induced with different combinations of aTc (100 ng/ml) and AHL (50 ng/ml) for 24 hours, after which the cultures were diluted in fresh media with or without IPTG (1 mM). These cultures were then incubated at 37°C for 8 hours and assayed for total LacZ activity with the FDG assay. (F) Cell populations that received both the input and write signals followed by the read signal exhibited enhanced levels of total LacZ activity. Error bars indicate SEM for three independent biological replicates.

of other inducible promoters (e.g., Fig. 3), could be written into genomic memory only when the write signal (Beta expression) is present. Thus, SCRIBE memory units can be readily reprogrammed, integrated with logic circuits, and decomposed into independent input, write, and read operations. We anticipate that more complex logic circuits could be combined with SCRIBE-based memory to create analog memory and computation systems capable of storing the results of multi-input calculations (32, 33).

## Discussion

We described a scalable platform that uses genomic DNA for analog, rewritable, and flexible memory distributed across living cell populations. One current limitation is the number of orthogonal inducible promoters that can be used as inputs, but this could be addressed by the development of additional inducible transcriptional regulatory devices (34). Additionally, ssDNA expression can be coupled to endogenous promoters to sense and record native cellular events. Although we primarily targeted mutations into functional genes to facilitate convenient functional and reporter assays, natural or synthetic noncoding DNA segments could also be used to record memory within genomic DNA. The recorded memory could then be read by high-throughput sequencing (fig. S2). A potential benefit of using synthetic DNA segments as memory registers is the ability to introduce mutations for memory storage that are neutral in terms of fitness costs.

SCRIBE enables conditional increases in the recombination rate at specific loci beyond background levels. The maximum observed recombination rate of the current SCRIBE platform ( $\sim 10^{-4}$  recombination events per generation) is suitable for long-term recording of analog memory distributed across the collective genomes of cellular populations (fig. S3). However, it is not high enough to allow recording of digital information and efficient genome editing at the single-cell level. In principle, population-level analog memory could be achieved by other types of DNA memory switches, such as site-specific recombinases, if they were tuned to achieve intermediate recombination rates. Further investigation is required to determine the exact mechanisms involved in processing retron-based ssDNAs for recombination into genomic DNA and the effects of different growth conditions on SCRIBE memory. Because Beta-mediated recombination is replication-dependent (27–29) and ssDNA is believed to be recombined into the genome during passage of the replication fork (27), we speculate that only actively dividing cells are likely to participate in the described population-level memory. Future optimization of SCRIBE [e.g., by modulating the mismatch repair system (14) and cellular exonucleases (35)] could lead to more efficient single-cell digital memories. This could enable other useful applications, including recording extracellular and intracellular events at the single-cell level for biological studies, dynamic engineering of cellular phenotypes, experimental evolution and population dynamics studies,

single-cell computation and memory, the construction of complex cellular state machines and biological Turing machines, and enhanced genome engineering techniques.

Additionally, because retrons have been found in a diverse range of microorganisms (20), in vivo ssDNA expression could potentially be extended to hard-to-transform organisms in which SCRIBE plasmids could be introduced by conjugation or transduction. Because retrons have also been shown to be functional in eukaryotes (24, 36, 37), they could potentially be used with other genome editing tools for memory. Moreover, by using error-prone RNA polymerases (38) and reverse transcriptases (39, 40), we anticipate that mutagenized ssDNA libraries could be generated inside cells for in vivo continuous evolution (41) and cellular barcoding applications. Finally, in vivo ssDNA generation could potentially be used to create DNA nanosystems (42–48) and ssDNA-protein hybrid nanomachines in living cells (49) or could be optimized and scaled-up to create an economical source of ssDNAs for DNA nanotechnology (50). In summary, we envision that in vivo ssDNA production and SCRIBE platforms will open up a broad range of new capabilities for engineering biology.

## Materials and methods

### Strains and plasmids

Conventional cloning methods were used to construct the plasmids. Lists of strains and plasmids used in this study and the construction procedures are provided in tables S1 and S2, respectively. The sequences for the synthetic parts and primers are provided in tables S3 and S4.

### Cells and antibiotics

Chemically competent *E. coli* DH5 $\alpha$  was used for cloning. Unless otherwise noted, antibiotics were used at the following concentrations: carbenicillin (50  $\mu$ g/ml), kanamycin (20  $\mu$ g/ml), chloramphenicol (30  $\mu$ g/ml), and spectinomycin (100  $\mu$ g/ml). In the experiment shown in Fig. 2, kanamycin (15  $\mu$ g/ml) and chloramphenicol (15  $\mu$ g/ml) were used.

### Detection of single-stranded DNA

Overnight cultures harboring IPTG-inducible plasmids encoding msd(wt), msd(wt) with deactivated RT [msd(wt)\_dRT], or msd(*kanR*)<sub>ON</sub> were grown overnight with or without IPTG (1 mM). Total RNA samples were prepared from non-induced or induced cultures using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. Total RNA (10  $\mu$ g) from each sample was treated with ribonuclease A (37°C, 2 hours) to remove RNA species and the msr moiety. The samples were then resolved on 10% tris-borate EDTA-urea denaturing gel and visualized with SYBR-Gold. A polyacrylamide gel electrophoresis-purified synthetic oligo (FF\_oligo347, 50 pmol) with the same sequence as ssDNA(wt) was used as a molecular size marker. The band intensities were measured by Fiji software (51). The intensities were normalized to the intensity of the marker

oligo, and normalized intensities were used to calculate the amount of ssDNA in each sample.

### Induction of cells and plating assays

For each experiment, three transformants were separately inoculated in Luria broth (LB) media plus appropriate antibiotics and grown overnight [37°C, 700 revolutions per minute (RPM)] to obtain seed cultures. Unless otherwise noted, inductions were performed by diluting the seed cultures (1:1000) in 2 ml of prewarmed LB plus appropriate antibiotics with or without inducers followed by 24 hours incubation (30°C, 700 RPM). Aliquots of the samples were then serially diluted, and appropriate dilutions were plated on selective media to determine the number of recombinants and viable cells in each culture. For each sample, the recombinant frequency was reported as the mean of the ratio of recombinants to viable cells for three independent replicates.

In all experiments, the number of viable cells was determined by plating aliquots of cultures on LB-plus-spectinomycin plates. LB-plus-kanamycin plates were used to determine the number of recombinants in the *kanR* reversion assay. For the *galK* reversion assay (Fig. 2), the numbers of *galK*<sub>ON</sub> recombinants were determined by plating the cells on MOPS EZ rich-defined media (Teknova) plus galactose (0.2%). The numbers of *galK*<sub>OFF</sub> recombinants were determined by plating the cells on MOPS EZ rich-defined media plus glycerol (0.2%) plus 2-DOG (2%). For the experiment shown in Fig. 3, the numbers of *kanR*<sub>ON</sub> *galK*<sub>ON</sub> and *kanR*<sub>OFF</sub> *galK*<sub>OFF</sub> cells were determined by using LB-plus-kanamycin plates and MOPS EZ rich-defined media plus glycerol (0.2%), 2-DOG (2%), and D-biotin (0.01%), respectively. The numbers of *kanR*<sub>ON</sub> *galK*<sub>OFF</sub> cells were determined by plating the cells on MOPS EZ-rich defined media plus glycerol (0.2%), 2-DOG (2%), kanamycin, and D-biotin (0.01%).

For the light-inducible SCRIBE experiment (Fig. 4A), induction was performed with white light (using the built-in fluorescent lamp in a VWR 1585 shaker incubator). The “dark” condition was achieved by wrapping aluminum foil around the tubes. Growth of and sampling from these cultures were performed as described earlier.

### LacZ assay

Overnight seed cultures were diluted (1:1000) in prewarmed LB plus appropriate antibiotics and inducers [with different concentrations of aTc or without aTc (Fig. 5, A to C) and with all the four possible combinations of aTc (100 ng/ml) and AHL (50 ng/ml) (Fig. 5, D to F)] and incubated for 24 hours (30°C, 700 RPM). These cultures were then diluted (1:50) in prewarmed LB plus appropriate antibiotics with or without IPTG (1 mM) and incubated for 8 hours (37°C, 700 RPM). To measure LacZ activity, 60  $\mu$ l of each culture was mixed with 60  $\mu$ l of B-PER II reagent (Pierce Biotechnology) and fluorescein di- $\beta$ -D-galactopyranoside (FDG) (0.05 mg/ml final concentration). The fluorescence signal (absorption/emission: 485/515) was monitored in a plate reader with continuous shaking for 2 hours. The LacZ activity was calculated by



normalizing the rate of FDG hydrolysis (obtained from fluorescence signal) to the initial optical density. For each sample, LacZ activity was reported as the mean of three independent biological replicates.

## REFERENCES AND NOTES

- N. Goldman *et al.*, Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **494**, 77–80 (2013). doi: [10.1038/nature11875](#); pmid: [23354052](#)
- G. M. Church, Y. Gao, S. Kosuri, Next-generation digital information storage in DNA. *Science* **337**, 1628 (2012). doi: [10.1126/science.1226355](#); pmid: [22903519](#)
- M. C. Inniss, P. A. Silver, Building synthetic memory. *Curr. Biol.* **23**, R812–R816 (2013). doi: [10.1016/j.cub.2013.06.047](#); pmid: [24028965](#)
- T. S. Gardner, C. R. Cantor, J. J. Collins, Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000). doi: [10.1038/35002131](#); pmid: [10659857](#)
- J. W. Kotula *et al.*, Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4838–4843 (2014). doi: [10.1073/pnas.1321321111](#); pmid: [24639514](#)
- D. R. Burrill, M. C. Inniss, P. M. Boyle, P. A. Silver, Synthetic memory circuits for tracking human cell fate. *Genes Dev.* **26**, 1486–1497 (2012). doi: [10.1101/gad.189035.112](#); pmid: [22751502](#)
- B. P. Kramer *et al.*, An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.* **22**, 867–870 (2004). doi: [10.1038/nbt980](#); pmid: [15184906](#)
- C. M. Ajo-Franklin *et al.*, Rational design of memory in eukaryotic cells. *Genes Dev.* **21**, 2271–2276 (2007). doi: [10.1101/gad.1586107](#); pmid: [17875664](#)
- P. Siuti, J. Yazbek, T. K. Lu, Synthetic circuits integrating logic and memory in living cells. *Nat. Biotechnol.* **31**, 448–452 (2013). doi: [10.1038/nbt.2510](#); pmid: [23396014](#)
- J. Bonnet, P. Subsoontorn, D. Endy, Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8884–8889 (2012). doi: [10.1073/pnas.1202344109](#); pmid: [22615351](#)
- J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, D. Endy, Amplifying genetic logic gates. *Science* **340**, 599–603 (2013). doi: [10.1126/science.1232758](#); pmid: [23539178](#)
- A. E. Friedland *et al.*, Synthetic gene networks that count. *Science* **324**, 1199–1202 (2009). doi: [10.1126/science.1172005](#); pmid: [19478183](#)
- H. M. Ellis, D. Yu, T. DiTizio, D. L. Court, High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6742–6746 (2001). doi: [10.1073/pnas.121164898](#); pmid: [11381128](#)
- N. Costantino, D. L. Court, Enhanced levels of  $\lambda$  Red-mediated recombinants in mismatch repair mutants. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15748–15753 (2003). doi: [10.1073/pnas.2434959100](#); pmid: [14673109](#)
- D. Yu *et al.*, An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5978–5983 (2000). doi: [10.1073/pnas.100127597](#); pmid: [10811905](#)
- J. A. Sawitzke *et al.*, Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. *J. Mol. Biol.* **407**, 45–59 (2011). doi: [10.1016/j.jmb.2011.01.030](#); pmid: [21256136](#)
- B. Swingle *et al.*, Oligonucleotide recombination in Gram-negative bacteria. *Mol. Microbiol.* **75**, 138–148 (2010). doi: [10.1111/j.1365-2958.2009.06976.x](#); pmid: [19943907](#)
- S. Datta, N. Costantino, X. Zhou, D. L. Court, Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1626–1631 (2008). doi: [10.1073/pnas.0709089105](#); pmid: [18230724](#)
- T. Yee, T. Furuichi, S. Inouye, M. Inouye, Multicopy single-stranded DNA isolated from a gram-negative bacterium, *Mycococcus xanthus*. *Cell* **38**, 203–209 (1984). doi: [10.1016/0092-8674\(84\)90541-5](#); pmid: [6088065](#)
- B. C. Lampson, M. Inouye, S. Inouye, Retrons, msDNA, and the bacterial genome. *Cytogenet. Genome Res.* **110**, 491–499 (2005). doi: [10.1159/000084982](#); pmid: [16093702](#)
- D. Lim, W. K. Maas, Reverse transcriptase-dependent synthesis of a covalently linked, branched DNA-RNA compound in *E. coli* B. *Cell* **56**, 891–904 (1989). doi: [10.1016/0092-8674\(89\)90693-4](#); pmid: [2466573](#)
- R. Lutz, H. Bujard, Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I<sub>1</sub>-I<sub>2</sub> regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997). doi: [10.1093/nar/25.6.1203](#); pmid: [9092630](#)
- P. L. Sharma, V. Nurpeisov, R. F. Schinazi, Retrovirus reverse transcriptases containing a modified YXDD motif. *Antivir. Chem. Chemother.* **16**, 169–182 (2005).pmid: [16004080](#)
- J. R. Mao, M. Shimada, S. Inouye, M. Inouye, Gene regulation by antisense DNA produced in vivo. *J. Biol. Chem.* **270**, 19684–19687 (1995). doi: [10.1074/jbc.270.34.19684](#); pmid: [7544343](#)
- S. Warming, N. Costantino, D. L. Court, N. A. Jenkins, N. G. Copeland, Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.* **33**, e36 (2005). doi: [10.1093/nar/gni035](#); pmid: [15731329](#)
- R. Ohlendörff, R. R. Vidavski, A. Eldar, K. Moffat, A. Möglich, From dusk till dawn: One-plasmid systems for light-regulated gene expression. *J. Mol. Biol.* **416**, 534–542 (2012). doi: [10.1016/j.jmb.2012.01.001](#); pmid: [22245580](#)
- M. S. Huen *et al.*, The involvement of replication in single stranded oligonucleotide-mediated gene repair. *Nucleic Acids Res.* **34**, 6183–6194 (2006). doi: [10.1093/nar/gkl852](#); pmid: [17088285](#)
- A. R. Poteete, Involvement of DNA replication in phage lambda Red-mediated homologous recombination. *Mol. Microbiol.* **68**, 66–74 (2008). doi: [10.1111/j.1365-2958.2008.06133.x](#); pmid: [18333884](#)
- A. R. Poteete, Involvement of *Escherichia coli* DNA replication proteins in phage Lambda Red-mediated homologous recombination. *PLOS ONE* **8**, e67440 (2013). doi: [10.1371/journal.pone.0067440](#); pmid: [23840702](#)
- B. E. Dutra, V. A. Suter Jr., S. T. Lovett, RecA-independent recombination is efficient but limited by exonucleases. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 216–221 (2007). doi: [10.1073/pnas.0608293104](#); pmid: [17182742](#)
- S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, R. Weiss, A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130–1134 (2005). doi: [10.1038/nature03461](#); pmid: [15858574](#)
- S. Ausländer, D. Ausländer, M. Müller, M. Wieland, M. Fussenegger, Programmable single-cell mammalian biocomputers. *Nature* **487**, 123–127 (2012).pmid: [22722847](#)
- R. Daniel, J. R. Rubens, R. Sarpeshkar, T. K. Lu, Synthetic analog computation in living cells. *Nature* **497**, 619–623 (2013). doi: [10.1038/nature12148](#); pmid: [23676681](#)
- E. J. Olson, L. A. Hartsough, B. P. Landry, R. Shroff, J. J. Tabor, Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. *Nat. Methods* **11**, 449–455 (2014). doi: [10.1038/nmeth.2884](#); pmid: [24608181](#)
- J. A. Mosberg, C. J. Gregg, M. J. Lajoie, H. H. Wang, G. M. Church, Improving lambda red genome engineering in *Escherichia coli* via rational removal of endogenous nucleases. *PLOS ONE* **7**, e44638 (2012). doi: [10.1371/journal.pone.0044638](#); pmid: [22957093](#)
- S. Miyata, A. Ohshima, S. Inouye, M. Inouye, In vivo production of a stable single-stranded cDNA in *Saccharomyces cerevisiae* by means of a bacterial retron. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5735–5739 (1992). doi: [10.1073/pnas.89.13.5735](#); pmid: [1378616](#)
- O. Mirochnitchenko, S. Inouye, M. Inouye, Production of single-stranded DNA in mammalian cells by means of a bacterial retron. *J. Biol. Chem.* **269**, 2380–2383 (1994).pmid: [7507924](#)
- S. Brakmann, S. Grzeszik, An error-prone T7 RNA polymerase mutant generated by directed evolution. *ChemBioChem* **2**, 212–219 (2001). doi: [10.1002/1439-7633\(20010302\)2:3<212::AID-CBIC212>3.0.CO;2-R](#); pmid: [11828447](#)
- J. D. Roberts, K. Bebenek, T. A. Kunkel, The accuracy of reverse transcriptase from HIV-1. *Science* **242**, 1171–1173 (1988). doi: [10.1126/science.2460925](#); pmid: [2460925](#)
- K. Bebenek, J. Abbotts, S. H. Wilson, T. A. Kunkel, Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots. *J. Biol. Chem.* **268**, 10324–10334 (1993).pmid: [7683675](#)
- K. M. Esvelt, J. C. Carlson, D. R. Liu, A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499–503 (2011). doi: [10.1038/nature09929](#); pmid: [21478873](#)
- Y. Amir *et al.*, Universal computing by DNA origami robots in a living animal. *Nat. Nanotechnol.* **9**, 353–357 (2014). doi: [10.1038/nnano.2014.58](#); pmid: [24705510](#)
- L. Qian, E. Winfree, J. Bruck, Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368–372 (2011). doi: [10.1038/nature10262](#); pmid: [21776082](#)
- G. Seelig, D. Soloveichik, D. Y. Zhang, E. Winfree, Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585–1588 (2006). doi: [10.1126/science.1132493](#); pmid: [17158324](#)
- P. W. Rothmund, Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006). doi: [10.1038/nature04586](#); pmid: [16541064](#)
- S. M. Douglas *et al.*, Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**, 414–418 (2009). doi: [10.1038/nature08016](#); pmid: [19458720](#)
- S. M. Douglas, I. Bachelet, G. M. Church, A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **335**, 831–834 (2012). doi: [10.1126/science.1214081](#); pmid: [22344439](#)
- S. M. Chirieleison, P. B. Allen, Z. B. Simpson, A. D. Ellington, X. Chen, Pattern transformation with DNA circuits. *Nat. Chem.* **5**, 1000–1005 (2013). doi: [10.1038/nchem.1764](#); pmid: [24256862](#)
- C. A. Brosey *et al.*, A new structural framework for integrating replication protein A into DNA processing machinery. *Nucleic Acids Res.* **41**, 2313–2327 (2013). doi: [10.1093/nar/gks1332](#); pmid: [23303776](#)
- S. Kosuri, G. M. Church, Large-scale de novo DNA synthesis: Technologies and applications. *Nat. Methods* **11**, 499–507 (2014). doi: [10.1038/nmeth.2918](#); pmid: [24781323](#)
- J. Schindelin *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi: [10.1038/nmeth.2019](#); pmid: [22743772](#)

## ACKNOWLEDGMENTS

This work was supported by the NIH New Innovator Award (1DP2OD008435), NIH National Centers for Systems Biology (1P50GM098792), the U.S. Office of Naval Research (N000141310424), and the Defense Advanced Research Projects Agency. Sequencing data have been deposited in GenBank with accession numbers KM923743 to KM923754.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/346/6211/1256272/suppl/DC1](http://www.sciencemag.org/content/346/6211/1256272/suppl/DC1)  
Supplementary Text  
Figs. S1 to S6  
Tables S1 to S4  
References (52–61)

19 May 2014; accepted 16 October 2014  
10.1126/science.1256272

## RESEARCH ARTICLE SUMMARY

## SYNTHETIC BIOLOGY

# Genomically encoded analog memory with precise in vivo DNA writing in living cell populations

Fahim Farzadfard and Timothy K. Lu\*

**INTRODUCTION:** The conversion of transient information into long-lasting responses is a common aspect of many biological processes and is crucial for the design of sophisticated synthetic circuits. Genomic DNA provides a rich medium for the storage of information in living cells. However, current cellular memory technologies are limited in their storage capacity and scalability.

**RATIONALE:** We converted genomic DNA into a “tape recorder” for memorizing information in living cell populations. This was achieved via SCRIBE (Synthetic Cellular Recorders Integrating Biological Events), a programmable and modular architecture for generating single-stranded DNA (ssDNA) inside of living cells in response to gene regulatory signals. When coexpressed with a recombinase, these ssDNAs address specific target loci on the basis of sequence homology and introduce precise mutations into genomic DNA, thus converting transient cellular signals into genomically encoded memory. This distributed biological memory leverages the large number of cells in bacterial cultures and encodes information into their collective genomic DNA in the form of the fraction of cells that carry specific mutations.

**RESULTS:** We show that SCRIBE enables the recording of arbitrary transcriptional inputs into DNA storage registers in living cells by translating regulatory signals into ssDNAs. In *Escherichia coli*, we expressed ssDNAs from engineered retransons that use a reverse transcriptase protein to produce hybrid RNA-ssDNA molecules. These intracellularly expressed ssDNAs are targeted

into specific genomic loci where they are recombined and converted into permanent memory. We show that genomically stored information can be readily reprogrammed by changing the ssDNA template and controlled via both chemical and light inputs. We demonstrate that genomically encoded memory can be read with a variety of techniques, including reporter genes, functional assays, and high-throughput DNA sequencing.

SCRIBE enables the recording of analog information such as the magnitude and time span of exposure to an input. This convenient feature is facilitated by the intermediate recombination rate of our current system ( $\sim 10^{-4}$  recombination events per generation), which we validated via a mathematical model and computer simulations. For example, we stored the overall exposure time to chemical inducers in the

DNA memory of bacterial populations for 12 days ( $\sim 120$  generations), independently of the induction pattern. The frequency of mutants in these populations was linearly related to the total exposure time.

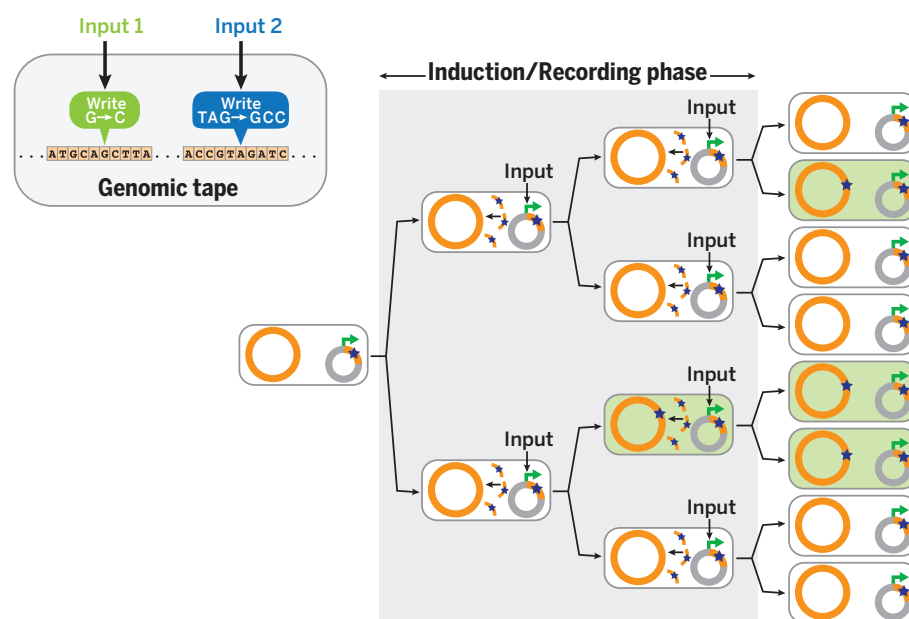
Furthermore, we demonstrate that SCRIBE-induced mutations can be written and erased and can be used to record multiple inputs across the distributed genomic DNA of bacterial populations. Finally, we show that SCRIBE memory can be decomposed into independent “input,” “write,” and “read” operations and used to create genetic “logic-and-memory” circuits, as well as “sample-and-hold” circuits.

**CONCLUSION:** We describe a scalable platform that uses genomic DNA for analog, rewritable, and flexible memory distributed across living cell populations. We anticipate that SCRIBE

## ON OUR WEB SITE

Read the full article at <http://dx.doi.org/10.1126/science.1256272>

will enable long-term cellular recorders for environmental and biomedical applications. Future optimization of recombination efficiencies achievable by SCRIBE could lead to more efficient single-cell digital memories and enhanced genome engineering technologies. Furthermore, the ability to regulate the generation of arbitrary targeted mutations with other gene-editing technologies should enable genomically encoded memory in additional organisms. ■



**SCRIBE enables distributed genomically encoded memory.** In the presence of an input, ssDNAs (orange curved lines) are produced from a plasmid-borne cassette (gray circles) and recombined into specific genomic loci (orange circles) that are targeted on the basis of sequence homology. This results in the accumulation of precise mutations (stars in green cells) as a function of the magnitude and duration of exposure to the input.

## RELATED ITEMS IN SCIENCE

S. Ausländer, M. Fussenegger, Dynamic genome engineering in living cells. *Science* **346**, 813–814 (2014).

The list of author affiliations is available in the full article online.

\*Corresponding author. E-mail: [timlu@mit.edu](mailto:timlu@mit.edu)

Cite this article as F. Farzadfard, T. K. Lu, *Science* **346**, 1256272 (2014). DOI: 10.1126/science.1256272



## RESEARCH ARTICLE

## SYNTHETIC BIOLOGY

# Genomically encoded analog memory with precise in vivo DNA writing in living cell populations

Fahim Farzadfard<sup>1,2,3</sup> and Timothy K. Lu<sup>1,2,3\*</sup>

Cellular memory is crucial to many natural biological processes and sophisticated synthetic biology applications. Existing cellular memories rely on epigenetic switches or recombinases, which are limited in scalability and recording capacity. In this work, we use the DNA of living cell populations as genomic “tape recorders” for the analog and distributed recording of long-term event histories. We describe a platform for generating single-stranded DNA (ssDNA) in vivo in response to arbitrary transcriptional signals. When coexpressed with a recombinase, these intracellularly expressed ssDNAs target specific genomic DNA addresses, resulting in precise mutations that accumulate in cell populations as a function of the magnitude and duration of the inputs. This platform could enable long-term cellular recorders for environmental and biomedical applications, biological state machines, and enhanced genome engineering strategies.

Due to its high storage capacity, durability, ease of duplication, and high-fidelity maintenance of information, DNA has garnered much interest as an artificial storage medium (1, 2). However, existing technologies for in vivo autonomous recording of information in cellular memory are limited in their storage capacity and scalability (3). Epigenetic memory devices such as bistable toggle switches (4–7) and positive-feedback loops (8) require orthogonal transcription factors and can lose their digital state due to environmental fluctuations or cell death. Recombinase-based memory devices enable the writing and storage of digital information in the DNA of living cells (9–12), where binary bits of information are stored in the orientation of large stretches of DNA. However, these devices do not efficiently exploit the full capacity of DNA for information storage: Recording a single bit of information with these devices often requires at least a few hundred base pairs of DNA, overexpression of a recombinase protein to invert the target DNA, and engineering recombinase recognition sites into target loci in advance. The scalability of this type of memory is further limited by the number of orthogonal recombinases that can be used in a single cell. Finally, epigenetic and recombinase-based memory devices described to date store digital information, and their recording capacity is exhausted within a

few hours of induction. Thus, these devices have not been adapted to record analog information, such as magnitude and time course of inputs over extended periods of time (i.e., multiple days or more).

Here we introduce SCRIBE (Synthetic Cellular Recorders Integrating Biological Events), a compact, modular strategy for producing single-stranded DNA (ssDNA) inside of living cells in response to a range of regulatory signals, such as small chemical inducers and light. These ssDNAs address specific target loci on the basis of sequence homology and introduce precise mutations into genomic DNA. The memory device can be easily reprogrammed to target different genomic locations by changing the ssDNA template. SCRIBE memory does not just record the absence or presence of arbitrary inputs (digital signals represented as binary “0s” or “1s”). Instead, by encoding information into the collective genomic DNA of cell populations, SCRIBE can track the magnitude and long-term temporal behavior of inputs, which are analog signals that can vary over a wide range of continuous values. This analog memory architecture leverages the large number of cells in bacterial cultures for distributed information storage and archives event histories in the fraction of cells in a population that carry specific mutations.

## Single-stranded DNA expression in living cells

Previously, it was shown that synthetic oligonucleotides delivered by electroporation into cells that overexpress Beta recombinase (from bacteriophage  $\lambda$ ) in *Escherichia coli* are specifically and efficiently recombined into homologous genomic sites (13–16). Thus, oligonucleotide-mediated recombination offers a powerful way to intro-

duce targeted mutations in a bacterial genome (17, 18). However, this technique requires the exogenous delivery of ssDNAs and cannot be used to couple arbitrary signals into genetic memory. To overcome these limitations, we developed a genome editing platform based on expressing ssDNAs inside of living cells by taking advantage of a widespread class of bacterial reverse transcriptases (RTs) called retrons (19, 20).

The wild-type (WT) retron cassette encodes three components in a single transcript: a RT protein and two RNA moieties, *msr* and *msd*, which act as the primer and the template for the reverse transcriptase, respectively (Fig. 1A, left). The *msr-msd* sequence in the retron cassette is flanked by two inverted repeats. Once transcribed, the *msr-msd* RNA folds into a secondary structure guided by the base pairing of the inverted repeats and the *msr-msd* sequence. The RT recognizes this secondary structure and uses a conserved guanosine residue in the *msr* as a priming site to reverse transcribe the *msd* sequence and produce a hybrid RNA-ssDNA molecule called msDNA (i.e., multicopy single-stranded DNA) (20, 21). To couple the expression of ssDNA to an external input, the WT Ec86 retron cassette from *E. coli* BL21 (21) was placed under the control of an isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG)-inducible promoter ( $P_{lacO}$ ) in *E. coli* DH5aPRO cells (22), which express high levels of the LacI and TetR repressors (Fig. 1A). The WT retron ssDNA [ssDNA(wt)] was readily detected in IPTG-induced cells, whereas no ssDNA was detected in noninduced cells (Fig. 1B). The identity of the detected ssDNA band was further confirmed by DNA sequencing (fig. S1). To verify that ssDNA expression depended on RT activity, point mutations [Asp<sup>197</sup>→Ala<sup>197</sup> (D197A) and D198A] were introduced to the active site of the RT to make a catalytically dead RT (dRT) (23). This modification completely abolished ssDNA production (Fig. 1B).

To engineer the *msd* template to express synthetic ssDNAs of interest, we initially tried to replace the whole *msd* sequence with a desired template. However, no ssDNA was detected, suggesting that some features of *msd* are required for ssDNA expression, as was previously noted for another retron (24). A variant in which the flanking regions of the *msd* stem remained intact (Fig. 1A, right) produced detectable amounts of ssDNA when induced by IPTG (Fig. 1B,  $P_{lacO\_msd}(kanR)_{ON}$  + IPTG). The correct identity of the detected ssDNA band was further confirmed by DNA sequencing (fig. S1). Thus, the lower part of the *msd* stem is essential for reverse transcription, whereas the upper part of the stem and the loop are dispensable and can be replaced with desired templates to produce ssDNAs of interest in vivo.

## Regulated genome editing with in vivo ssDNAs

To demonstrate that intracellularly expressed ssDNAs can be recombined into target genomic loci by concomitant expression of Beta, we developed a selectable marker reversion assay (Fig. 1C). The *kanR* gene, which encodes neomycin

<sup>1</sup>Synthetic Biology Group, Research Laboratory of Electronics, Department of Electrical Engineering and Computer Science and Department of Biological Engineering, Massachusetts Institute of Technology (MIT), 77 Massachusetts Avenue, Cambridge, MA 02139, USA. <sup>2</sup>MIT Synthetic Biology Center, 500 Technology Square, Cambridge, MA 02139, USA. <sup>3</sup>MIT Microbiology Program, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. \*Corresponding author. E-mail: timlu@mit.edu

phosphotransferase II and confers resistance to kanamycin (Kan), was integrated into the *galK* locus. Two stop codons were then introduced into the genomic *kanR* to make a Kan-sensitive *kanR<sub>OFF</sub>* reporter strain (DH5aPRO *galK::kanR<sub>W28TAA, A29TAG</sub>*). These premature stop codons could be reverted back to the WT sequence via recombination with engineered ssDNA(*kanR*)<sub>ON</sub>, thus conferring kanamycin resistance (Fig. 1C). Specifically, ssDNA(*kanR*)<sub>ON</sub> contains 74 base pairs of homology to the regions of the *kanR<sub>OFF</sub>* locus flanking the premature stop codons and replaces the stop codons with the WT *kanR* gene sequence (Fig. 1C).

We cloned the Beta gene (*bet*) into a plasmid under the control of the anhydrotetracycline (aTc)-inducible *P<sub>tetO</sub>* promoter and introduced it along with the IPTG-inducible *msd(kanR)*<sub>ON</sub> construct into the *kanR<sub>OFF</sub>* strain (Fig. 1C). Induction of cultures harboring these two plasmids with either IPTG (1 mM) or aTc (100 ng/ml) resulted in a slight increase in the frequency of Kan-resistant cells within the population (Fig. 1C). However, coexpression of both ssDNA(*kanR*)<sub>ON</sub> and Beta with IPTG and aTc resulted in a  $>10^4$ -fold increase in the recombinant frequency relative to the non-induced cells. This corresponded to a  $>10^3$ -fold

increase relative to cells induced with IPTG only and a 60-fold increase relative to cells induced with aTc only. This increase in the recombinant frequency was dependent on the RT activity, as it was largely abolished with dRT. The genotypes of randomly selected Kan-resistant colonies were further confirmed by DNA sequencing to contain precise reversions of the two codons to the WT sequence (fig. S1). No Kan-resistant colonies were detected when a nonspecific ssDNA [ssDNA(wt)] was coexpressed with Beta in the *kanR<sub>OFF</sub>* reporter cells, confirming that Kan-resistant cells were not produced due to spontaneous mutations. In additional experiments, we used high-throughput sequencing (Illumina HiSeq) on the bacterial populations to analyze the genomically encoded memory (see supplementary materials and fig. S2). Comparable recombinant frequencies were obtained from both the plating assay and sequencing, confirming that genomically encoded memory can be read without the need for functional assays and reporters.

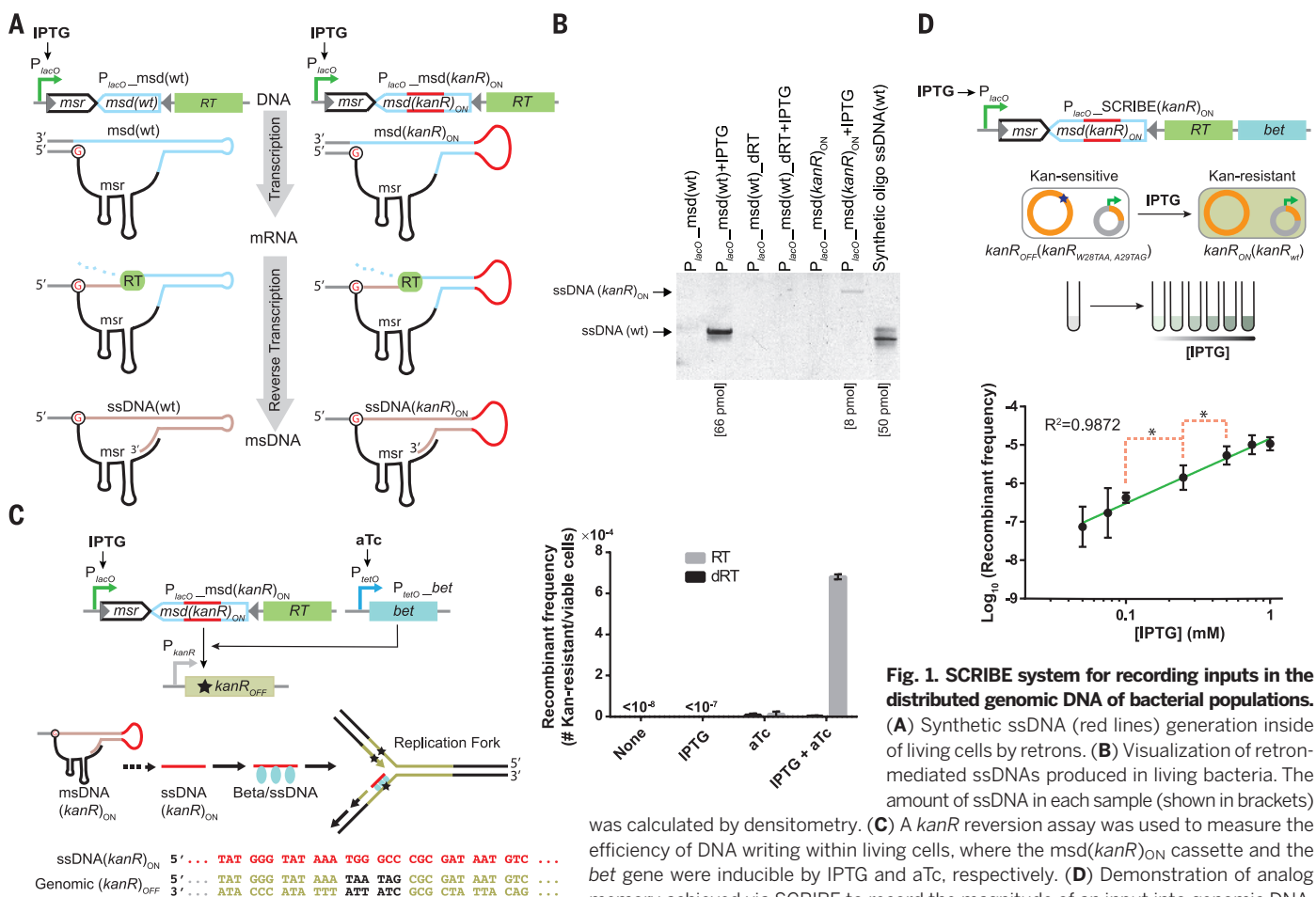
### Recording input magnitudes into genomic memory

We reasoned that the rate of recombination between engineered ssDNAs and genomic DNA

could be effectively modulated by changing expression levels of the engineered retron cassette and Beta. This feature would enable the recording of analog information, such as the magnitude of an input signal, in the proportion of cells in a population with a specific mutation in genomic DNA. To demonstrate this, both the ssDNA(*kanR*)<sub>ON</sub> expression cassette and *bet* were placed into a single synthetic operon [hereafter referred to as the SCRIBE(*kanR*)<sub>ON</sub> cassette] under the control of *P<sub>lacO</sub>* (Fig. 1D). The *kanR<sub>OFF</sub>* reporter cells harboring this synthetic operon were induced with different concentrations of IPTG. The fraction of Kan-resistant recombinants increased linearly with the input inducer concentration on a log-log plot over a range of  $\sim 10^{-7}$  to  $\sim 10^{-5}$  (Fig. 1D). Statistical tests showed that at least four different concentrations of the inducer (including 0 mM IPTG) could be resolved in this experiment. Thus, the efficiency of genome writing in a population can be quantitatively tuned with external inputs.

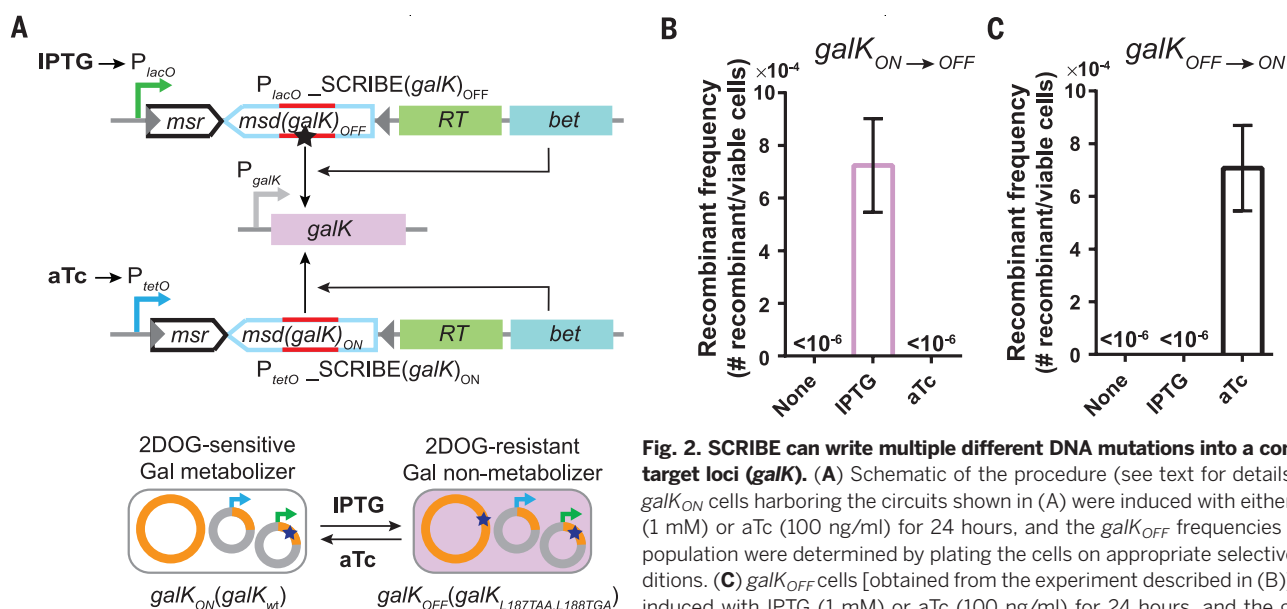
### Writing and rewriting genomic memory

We next created a complementary set of SCRIBE cassettes to write and erase (rewrite) information in the genomic *galK* locus using two different chemical inducers. Cells expressing *galK* can



The green line is a linear regression fit. The red dashed brackets marked with asterisks connect the closest data points that are statistically significant with respect to each other ( $P < 0.05$  based on one-tailed Welch's *t* test). Error bars indicate SEM for three independent biological replicates.





**Fig. 2. SCRIPE can write multiple different DNA mutations into a common target loci (*galk*).** (A) Schematic of the procedure (see text for details). (B) *galk<sub>ON</sub>* cells harboring the circuits shown in (A) were induced with either IPTG (1 mM) or aTc (100 ng/ml) for 24 hours, and the *galk<sub>OFF</sub>* frequencies in the population were determined by plating the cells on appropriate selective conditions. (C) *galk<sub>OFF</sub>* cells [obtained from the experiment described in (B)] were induced with IPTG (1 mM) or aTc (100 ng/ml) for 24 hours, and the *galk<sub>ON</sub>* frequencies in the population were determined by plating the cells on appropriate selective conditions. Error bars indicate SEM for three independent biological replicates.

metabolize and grow on galactose as the sole carbon source. However, these *galk<sub>ON</sub>* cells cannot metabolize 2-deoxy-galactose (2DOG) and cannot grow on plates containing glycerol (carbon source) + 2DOG. On the other hand, *galk*-negative (*galk<sub>OFF</sub>*) cells cannot grow on galactose as the sole carbon source but can grow on glycerol + 2DOG plates (25). We transformed DH5αPRO *galk<sub>ON</sub>* cells with plasmids encoding IPTG-inducible SCRIPE(*galk<sub>OFF</sub>*) and aTc-inducible SCRIPE(*galk<sub>ON</sub>*) cassettes (Fig. 2A). Induction of SCRIPE(*galk<sub>OFF</sub>*) by IPTG resulted in the writing of two stop codons into *galk<sub>ON</sub>*, leading to *galk<sub>OFF</sub>* cells that could grow on glycerol + 2DOG plates (Fig. 2B). Induction of SCRIPE(*galk<sub>ON</sub>*) in these *galk<sub>OFF</sub>* cells with aTc reversed the IPTG-induced modification, leading to *galk<sub>ON</sub>* cells that could grow on galactose plates (Fig. 2C). These results show that writing on genomic DNA with SCRIPE is reversible and that distinct information can be written and rewritten into the same locus.

### Writing multiple mutations into independent loci

Scaling the capacity of previous memory devices is challenging because each additional bit of information requires additional orthogonal proteins (e.g., recombinases or transcription factors). In contrast, orthogonal SCRIPE memory devices are potentially easier to scale because they can be built by simply changing the ssDNA template (*msd*). To demonstrate this, we used SCRIPE to record multiple independent inputs into different genomic loci of bacterial population. We integrated the *kanR<sub>OFF</sub>* reporter gene into the *bioA* locus of DH5αPRO to create a *kanR<sub>OFF</sub>galk<sub>ON</sub>* strain. These cells were then transformed with plasmids encoding IPTG-inducible SCRIPE(*kanR<sub>ON</sub>*) and aTc-inducible SCRIPE(*galk<sub>OFF</sub>*) cassettes (Fig. 3A). Induction of these cells with IPTG or aTc resulted

in the production of cells with phenotypes corresponding to *kanR<sub>ON</sub>galk<sub>ON</sub>* or *kanR<sub>OFF</sub>galk<sub>OFF</sub>* genotypes, respectively (Fig. 3, B and C). Comparable numbers of *kanR<sub>ON</sub>galk<sub>ON</sub>* and *kanR<sub>OFF</sub>galk<sub>OFF</sub>* cells ( $\sim 2 \times 10^{-4}$  and  $\sim 3 \times 10^{-4}$  recombinant/viable cells, respectively) were produced when the cultures were induced with both aTc and IPTG (Fig. 3C, left panel). Furthermore, very few individual colonies ( $\sim 3 \times 10^{-7}$  recombinant/viable cells) containing both writing events (*kanR<sub>ON</sub>galk<sub>OFF</sub>*) were obtained in the cultures that were induced with both aTc and IPTG (Fig. 3C, right panel). These data suggest that although multiplexed writing at single-cell level is rare with SCRIPE's current level of recombination efficiency, multiple independent inputs can be successfully recorded into the distributed genomic DNA of bacterial subpopulations.

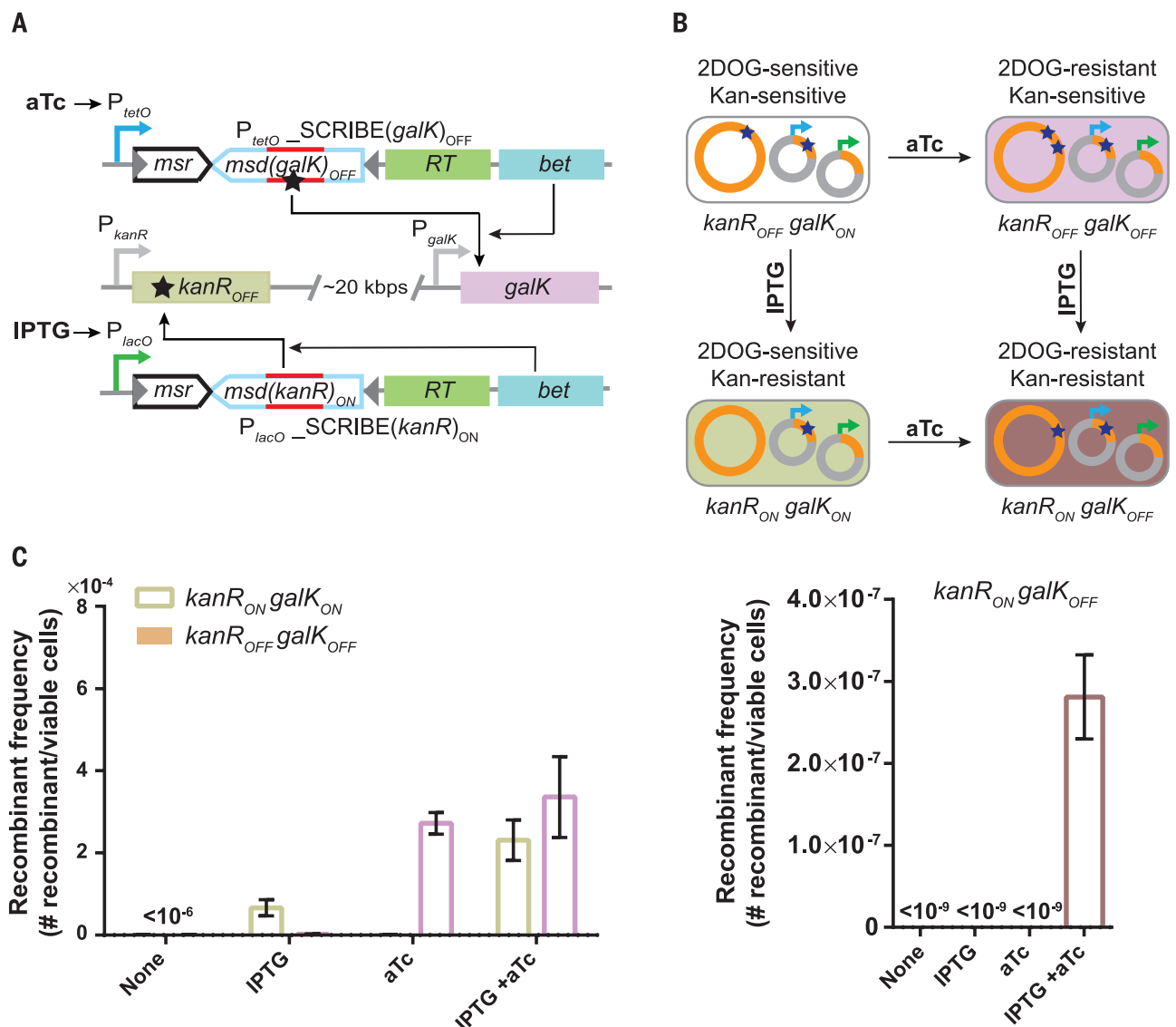
### Optogenetic genome editing for light-to-DNA memory

In SCRIPE, the expression of each individual ssDNA can be triggered by any endogenous or exogenous signal that can be coupled into transcriptional regulation, thus recording these inputs into long-lasting DNA storage. In addition to small-molecule chemicals, we showed that light can be used to trigger specific genome editing for genomically encoded memory. We placed the SCRIPE(*kanR<sub>ON</sub>*) cassette under the control of a previously described light-inducible promoter (*P<sub>Dawn</sub>*) (26) within *kanR<sub>OFF</sub>* cells (Fig. 4A). These cultures were then grown for 4 days in the presence of light or in the dark (Fig. 4A). As Beta-mediated recombination is reportedly replication-dependent (27–29), dilutions of these cultures were made into fresh media at the end of each day to maintain active replication in the cultures. At the end of each day, samples were taken to determine the number of Kan-resistant and viable cells (Fig. 4A). Cultures grown in the dark yielded undetectable

levels of Kan-resistant cells (Fig. 4A). In contrast, the frequency of Kan-resistant cells increased steadily over time in the cultures that were grown in the presence of light, indicating the successful recording of light input into long-lasting DNA memory. The analog memory faithfully stored the total time of light exposure, rather than just the digital presence or absence of light.

### Recording the exposure time of inputs

The linear increase in the frequency of Kan-resistant colonies over time due to exposure to light indicates that the duration of inputs can be recorded into population-wide DNA memory using SCRIPE. To further explore population-wide genomically encoded memory whose state is a function of input exposure time, we used the *kanR<sub>OFF</sub>* strain harboring the constructs shown in Fig. 1C, where expression of ssDNA(*kanR<sub>ON</sub>*) and Beta are controlled by IPTG and aTc, respectively. These cells were subjected to four different patterns of the inputs for 12 successive days (patterns I to IV, Fig. 4B). Kan-resistant cells did not accumulate in the negative control (pattern I), which was never exposed to the inducers. The fraction of Kan-resistant cells in the three other patterns (II, III, and IV) increased linearly over their respective induction periods and remained relatively constant when the inputs were removed. These data indicate that the genomically encoded memory was stable in the absence of the inputs over the course of the experiment. The recombinant frequencies in patterns III and IV, which were induced for the same total amount of time but with different temporal patterns, reached comparable levels at the end of the experiment. These data demonstrate that the genomic memory integrates over the total induction time and is independent of the input pattern, and therefore can be used to stably record event histories over many days.



**Fig. 3. Writing multiple mutations into independent target loci within population.** (A) Constructs used to target genomic *kanR*<sub>OFF</sub> and *galK*<sub>ON</sub> loci with IPTG-inducible and aTc-inducible SCRIBE cassettes, respectively. (B) Induction of *kanR*<sub>OFF</sub> *galK*<sub>ON</sub> cells with IPTG or aTc generates cells with the *kanR*<sub>ON</sub> *galK*<sub>ON</sub> or *kanR*<sub>OFF</sub> *galK*<sub>OFF</sub> genotypes, respectively. Induction of *kanR*<sub>OFF</sub> *galK*<sub>ON</sub> cells with both IPTG

and aTc generates cells with the *kanR*<sub>ON</sub> *galK*<sub>OFF</sub> genotype. (C) *kanR*<sub>OFF</sub> *galK*<sub>ON</sub> reporter cells containing the circuits in (A) were induced with different combinations of IPTG (1 mM) and aTc (100 ng/ml) for 24 hours at 30°C, and the fraction of cells with the various genotypes were determined by plating the cells on appropriate selective media. Error bars indicate SEM for three independent biological replicates.

The linear increase in the fraction of recombinants in the induced cell populations over time was consistent with a deterministic model (dashed lines in Fig. 4B, also see supplementary materials). Specifically, when triggered by inputs, SCRIBE can substantially increase the rate of recombination events at a specific target site above the WT rate [which is reportedly <10<sup>-10</sup> events per generation in *recA*<sup>-</sup> background (30)]. When recombination rates are ~10<sup>-4</sup> events per generation, which is consistent with the recombination rate estimated for SCRIBE from data in Fig. 4B, a simple deterministic model and a detailed stochastic simulation both predict a linear increase in the frequency of recombinant alleles in a population over time, as long as this frequency is less than a few percent and cells in the population are equally fit over the time scale of interest (see

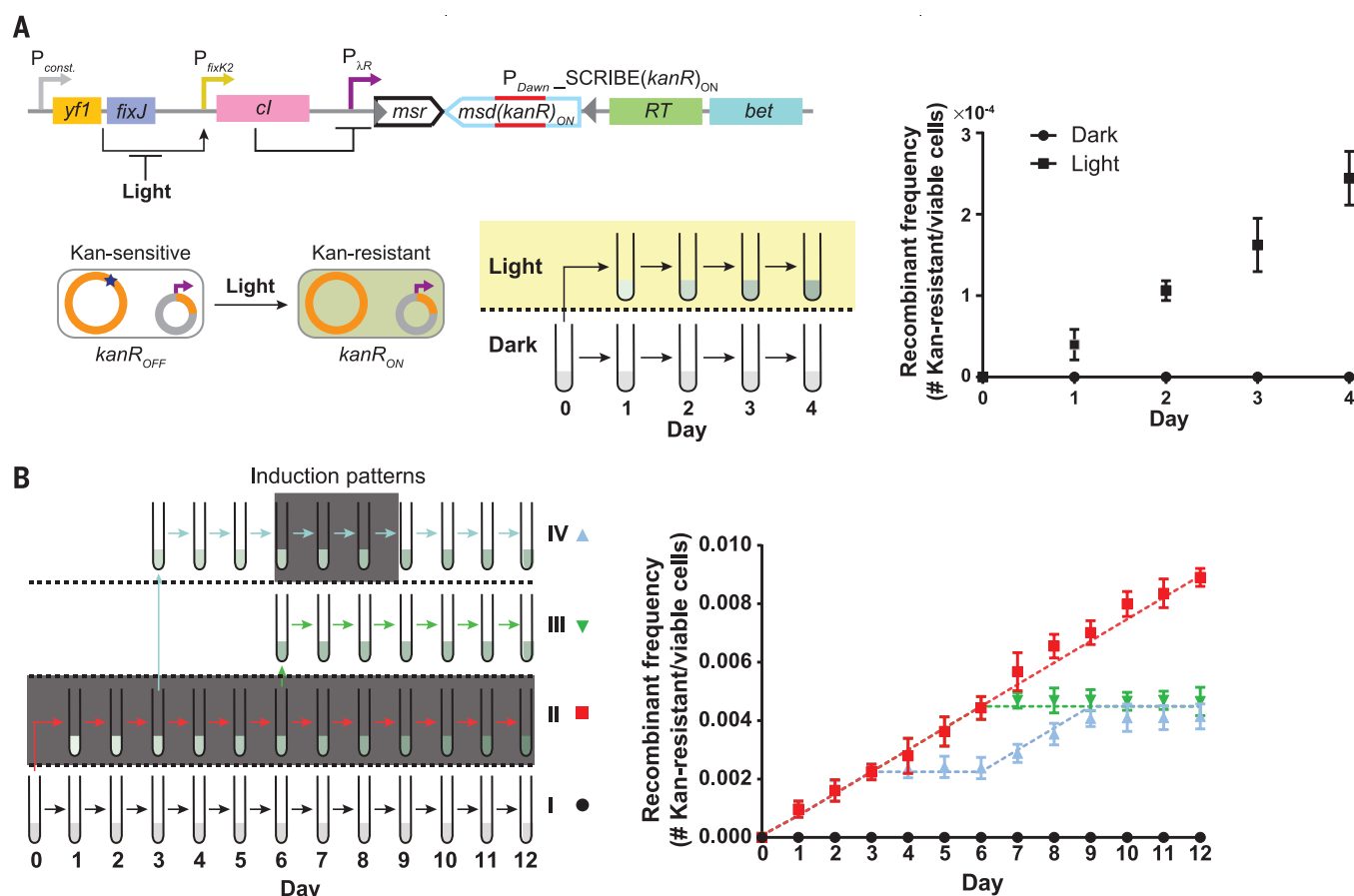
supplementary materials and figs. S3 and S4). These models enable one to determine the ideal range of recombination rates for a given application, which depends on parameters such as the frequency of dilution, the sensitivity of the method used for reading the memory, the desired input duration to be recorded, and so forth. For example, recombination rates that are too low would be challenging to quantify and could result in loss of memory if the cultures were diluted. Moreover, higher recombination rates lead to more rapid saturation of memory capacity in which the system is unable to provide a straightforward linear relation between the input exposure time and the state of the memory (fig. S3). Thus, intermediate levels of recombination rates are desirable for population-level analog memory units that can record the

time span of exposure to inputs (see supplementary materials).

### Decoupling memory operations

SCRIBE memory can be used to create more complex synthetic memory circuits. To demonstrate this, we first built a synthetic gene circuit that can record different input magnitudes into DNA memory. The memory state can then be read out later (after the initial input is removed) upon addition of a secondary signal. Specifically, we built an IPTG-inducible *lacZ*<sub>OFF</sub> (*lacZ*<sub>A35TAA, S36TAG</sub>) reporter construct in DH5αPRO cells (Fig. 5A). Expression of this reporter is normally repressed except when IPTG (“read” signal, Fig. 5A) is added, thus enabling a convenient and switchable population-level readout of the memory based on total LacZ activity (Fig. 5B). The *lacZ*<sub>OFF</sub>





**Fig. 4. Optogenetic genome editing and analog memory for long-term recording of input signal exposure times in the genomic DNA of living cell populations.** (A) We coupled expression of *SCRIBE(kanR)<sub>ON</sub>* to an optogenetic system (*P<sub>Dawn</sub>*). The *yf1/fixJ* synthetic operon was expressed from a constitutive promoter: In dark conditions, YF1 interacts with and phosphorylates FixJ. Phosphorylated FixJ activates the *P<sub>fixK2</sub>* promoter, which drives  $\lambda$  repressor (*cl*) expression, which subsequently represses the *SCRIBE(kanR)<sub>ON</sub>* cassette. Light inhibits the interaction between YF1 and FixJ, leading to the generation of ssDNA(*kanR*)<sub>ON</sub> and Beta expression and, thus, the conversion of *kanR<sub>OFF</sub>* to *kanR<sub>ON</sub>*. Cells harboring this circuit were grown overnight at 37°C in the dark, diluted 1:1000, and then incubated for 24 hours at 30°C in the dark (no shading) or in the presence of light (yellow shading). Subsequently, cells were diluted by 1:1000

and grown for another 24 hours at 30°C in the dark or in the presence of light. The dilution-regrowth cycle was performed for four consecutive days. The *kanR* allele frequencies in the populations were determined by sampling the cultures after each 24-hour period. (B) *SCRIBE* analog memory records the total time exposure to a given input, regardless of the underlying induction pattern. Cells harboring the circuit shown in Fig. 1C were grown in four different patterns (I to IV) over a 12-day period, where induction by IPTG (1 mM) and aTc (100 ng/mL) is represented by dark gray shading. At the end of each 24-hour incubation period, cells were diluted by 1:1000 into fresh media. The frequency of Kan-resistant cells in the cultures was determined at the end of each day. Dashed lines represent the recombinant allele frequencies predicted by the model (see supplementary materials). Error bars indicate SEM for three independent biological replicates.

reporter cells were transformed with a plasmid encoding an aTc-inducible *SCRIBE(lacZ)<sub>ON</sub>* cassette (Fig. 5A). Overnight cultures were diluted and induced with various amounts of aTc to write the genomic memory (Fig. 5B). These cells were grown up to saturation and then diluted into fresh media in the presence or absence of IPTG to read the genomic memory (Fig. 5B). In the absence of IPTG, the total LacZ activity remained low, regardless of the aTc concentration. In the presence of IPTG, cultures that had been exposed to higher aTc concentrations had greater total LacZ activity. These results show that population-level reading of genomically encoded memory can be decoupled from writing and controlled externally. Furthermore, this circuit enables the magnitude of the inducer (aTc) to be stably recorded in the distributed genomic memory of a cellular population. Independent control over the read memory operation

as shown in this experiment could help to minimize fitness costs associated with the expression of reporter genes until needed.

We have shown that (i) both ssDNA expression and Beta are required for writing into genomic memory (Fig. 1C), (ii) multiple ssDNAs can be used to independently address different memory units (Fig. 3), and (iii) genomic memory is stably recorded into DNA and can be used to modify functional genes whose expression can be controlled by external inducers (Figs. 1 to 4). Thus, *SCRIBE* memory units can be conceptually decomposed into separate “input,” “write,” and “read” operations to facilitate greater control and the integration of logic with memory. The separation of these signals could enable master control over the writing of multiple independent inputs into genomic memory. To achieve this, we placed the *msd(lacZ)<sub>ON</sub>* cassette under the control of an acyl homoserine lactone (AHL)-inducible promoter

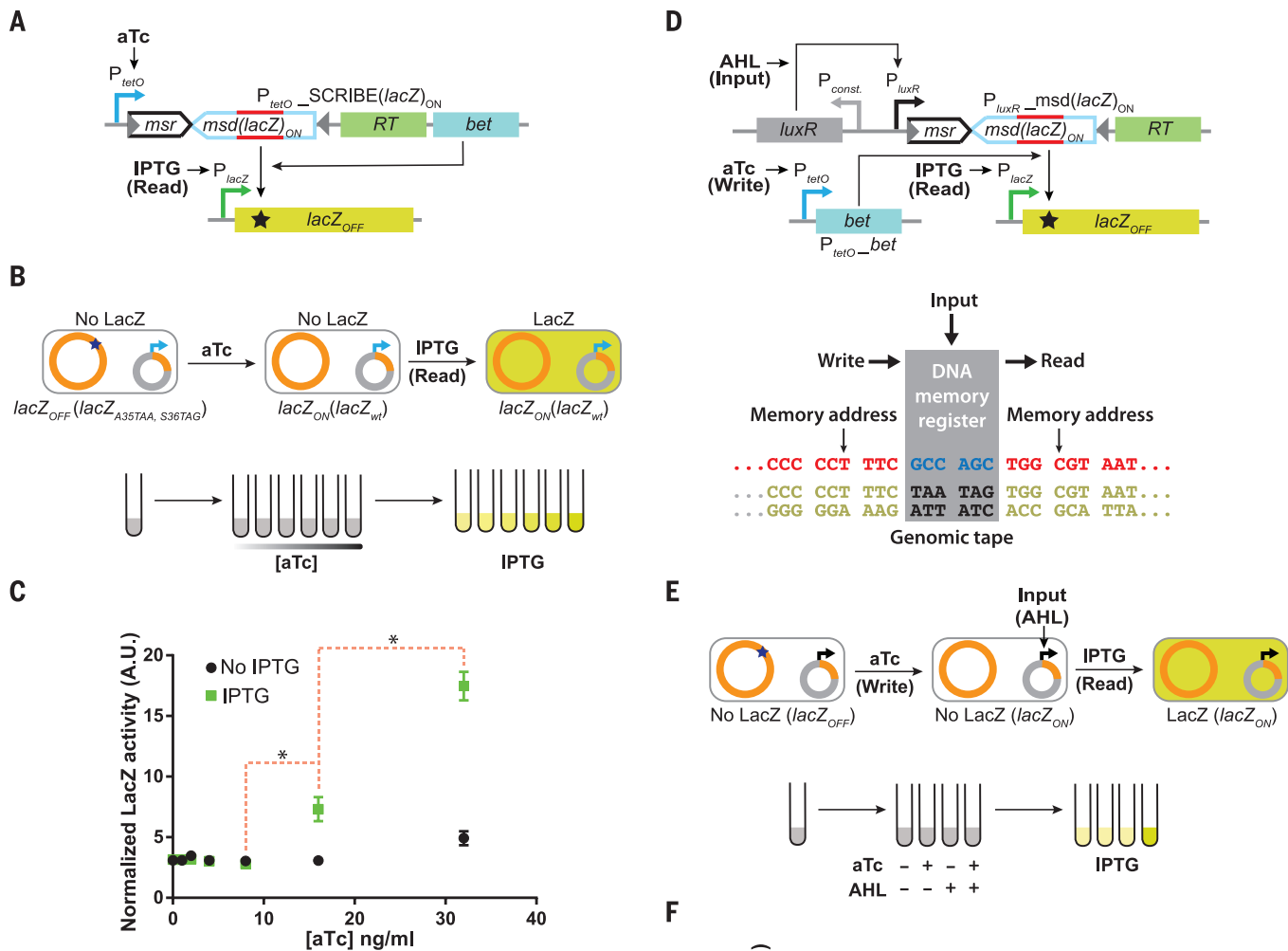
(*P<sub>luxR</sub>*) (31) and cotransformed this plasmid with an aTc-inducible Beta-expressing plasmid into the *lacZ<sub>OFF</sub>* reporter strain (Fig. 5D). Using this design, information on the input (ssDNA expression via addition of AHL) can be written into DNA memory only in the presence of the write signal (Beta expression via addition of aTc). The information recorded in the memory register (i.e., the state of *lacZ* across the population) can be retrieved by adding the read signal (IPTG).

To demonstrate this, overnight *lacZ<sub>OFF</sub>* cultures harboring the circuit shown in Fig. 5D were diluted and then grown to saturation in the presence of all four possible combinations of AHL and aTc (Fig. 5E). The saturated cultures were then diluted into fresh media in the absence or presence of IPTG. As shown in Fig. 5F, only cultures that had been exposed to both the input and write signals simultaneously showed substantial LacZ activity, and only when they were

induced with the read signal. These results indicate that short stretches of DNA of living organisms can be used as addressable read/write memory registers to record transcriptional inputs.

Furthermore, SCRIBE memory can be combined with logic, such as the AND function between the input and write signals shown here. The logic in Fig. 5D enables this circuit to act as a “sample-

and-hold” system in which information about an input can be recorded in the presence of another signal and read out at will. Additional inputs in the form of orthogonal ssDNAs under the control



**Fig. 5. SCRIBE memory operations can be decoupled into independent “input,” “write,” and “read” operations, thus facilitating greater control over addressable memory registers in genomic tape recorders and the creation of sample-and-hold**

**circuits. (A)** We built a circuit where information about the first inducer (aTc) is recorded in the population, which can then be read later upon addition of a second inducer (IPTG) that triggers a read operation. We created an IPTG-inducible *lacZ*<sub>OFF</sub> locus in the DH5αPRO background, which contains the full-length *lacZ* gene with two premature stop codons inside the open-reading frame. Expression of ssDNA(*lacZ*)<sub>ON</sub> from the aTc-inducible SCRIBE(*lacZ*)<sub>ON</sub> cassette results in the reversion of the stop codons inside *lacZ*<sub>OFF</sub> to yield the *lacZ*<sub>ON</sub> genotype. **(B)** Cells harboring the circuit shown in (A) were grown in the presence of different levels of aTc for 24 hours at 30°C to enable recording into genomic DNA. Subsequently, cell populations were diluted into fresh media without or with IPTG

(1 mM) and incubated at 37°C for 8 hours. **(C)** Total LacZ activity in these cultures was measured using a fluorogenic *lacZ* substrate (FDG) assay. The red dashed brackets marked with asterisks connect the closest data points of IPTG-induced samples that are statistically significant ( $P < 0.05$  based on one-tailed Welch's  $t$  test). A.U., arbitrary units. **(D)** We extended the circuit in (A) to create a sample-and-hold circuit where input, write, and read operations are independently controlled. This feature enables the creation of addressable read/write memory registers in the genomic DNA tape. Induction of cells with the input signal (AHL) produces ssDNA(*lacZ*)<sub>ON</sub>, which targets the genomic *lacZ*<sub>OFF</sub> locus for reversion to the WT sequence. In the presence of the write signal (aTc), which expresses Beta, ssDNA(*lacZ*)<sub>ON</sub> is recombined into the *lacZ*<sub>OFF</sub> locus and produces the *lacZ*<sub>ON</sub> genotype. Thus, the write signal enables the input signal to be sampled and held in memory. The total LacZ activity in the cell populations is retrieved by adding the read signal (IPTG). **(E)** Cells harboring the circuit shown in (D) were induced with different combinations of aTc (100 ng/ml) and AHL (50 ng/ml) for 24 hours, after which the cultures were diluted in fresh media with or without IPTG (1 mM). These cultures were then incubated at 37°C for 8 hours and assayed for total LacZ activity with the FDG assay. **(F)** Cell populations that received both the input and write signals followed by the read signal exhibited enhanced levels of total LacZ activity. Error bars indicate SEM for three independent biological replicates.



of other inducible promoters (e.g., Fig. 3), could be written into genomic memory only when the write signal (Beta expression) is present. Thus, SCRIBE memory units can be readily reprogrammed, integrated with logic circuits, and decomposed into independent input, write, and read operations. We anticipate that more complex logic circuits could be combined with SCRIBE-based memory to create analog memory and computation systems capable of storing the results of multi-input calculations (32, 33).

## Discussion

We described a scalable platform that uses genomic DNA for analog, rewritable, and flexible memory distributed across living cell populations. One current limitation is the number of orthogonal inducible promoters that can be used as inputs, but this could be addressed by the development of additional inducible transcriptional regulatory devices (34). Additionally, ssDNA expression can be coupled to endogenous promoters to sense and record native cellular events. Although we primarily targeted mutations into functional genes to facilitate convenient functional and reporter assays, natural or synthetic noncoding DNA segments could also be used to record memory within genomic DNA. The recorded memory could then be read by high-throughput sequencing (fig. S2). A potential benefit of using synthetic DNA segments as memory registers is the ability to introduce mutations for memory storage that are neutral in terms of fitness costs.

SCRIBE enables conditional increases in the recombination rate at specific loci beyond background levels. The maximum observed recombination rate of the current SCRIBE platform ( $\sim 10^{-4}$  recombination events per generation) is suitable for long-term recording of analog memory distributed across the collective genomes of cellular populations (fig. S3). However, it is not high enough to allow recording of digital information and efficient genome editing at the single-cell level. In principle, population-level analog memory could be achieved by other types of DNA memory switches, such as site-specific recombinases, if they were tuned to achieve intermediate recombination rates. Further investigation is required to determine the exact mechanisms involved in processing retron-based ssDNAs for recombination into genomic DNA and the effects of different growth conditions on SCRIBE memory. Because Beta-mediated recombination is replication-dependent (27–29) and ssDNA is believed to be recombined into the genome during passage of the replication fork (27), we speculate that only actively dividing cells are likely to participate in the described population-level memory. Future optimization of SCRIBE [e.g., by modulating the mismatch repair system (14) and cellular exonucleases (35)] could lead to more efficient single-cell digital memories. This could enable other useful applications, including recording extracellular and intracellular events at the single-cell level for biological studies, dynamic engineering of cellular phenotypes, experimental evolution and population dynamics studies,

single-cell computation and memory, the construction of complex cellular state machines and biological Turing machines, and enhanced genome engineering techniques.

Additionally, because retrons have been found in a diverse range of microorganisms (20), in vivo ssDNA expression could potentially be extended to hard-to-transform organisms in which SCRIBE plasmids could be introduced by conjugation or transduction. Because retrons have also been shown to be functional in eukaryotes (24, 36, 37), they could potentially be used with other genome editing tools for memory. Moreover, by using error-prone RNA polymerases (38) and reverse transcriptases (39, 40), we anticipate that mutagenized ssDNA libraries could be generated inside cells for in vivo continuous evolution (41) and cellular barcoding applications. Finally, in vivo ssDNA generation could potentially be used to create DNA nanosystems (42–48) and ssDNA-protein hybrid nanomachines in living cells (49) or could be optimized and scaled-up to create an economical source of ssDNAs for DNA nanotechnology (50). In summary, we envision that in vivo ssDNA production and SCRIBE platforms will open up a broad range of new capabilities for engineering biology.

## Materials and methods

### Strains and plasmids

Conventional cloning methods were used to construct the plasmids. Lists of strains and plasmids used in this study and the construction procedures are provided in tables S1 and S2, respectively. The sequences for the synthetic parts and primers are provided in tables S3 and S4.

### Cells and antibiotics

Chemically competent *E. coli* DH5 $\alpha$  was used for cloning. Unless otherwise noted, antibiotics were used at the following concentrations: carbenicillin (50  $\mu$ g/ml), kanamycin (20  $\mu$ g/ml), chloramphenicol (30  $\mu$ g/ml), and spectinomycin (100  $\mu$ g/ml). In the experiment shown in Fig. 2, kanamycin (15  $\mu$ g/ml) and chloramphenicol (15  $\mu$ g/ml) were used.

### Detection of single-stranded DNA

Overnight cultures harboring IPTG-inducible plasmids encoding msd(wt), msd(wt) with deactivated RT [msd(wt)\_dRT], or msd(*kanR*)<sub>ON</sub> were grown overnight with or without IPTG (1 mM). Total RNA samples were prepared from non-induced or induced cultures using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. Total RNA (10  $\mu$ g) from each sample was treated with ribonuclease A (37°C, 2 hours) to remove RNA species and the msr moiety. The samples were then resolved on 10% tris-borate EDTA-urea denaturing gel and visualized with SYBR-Gold. A polyacrylamide gel electrophoresis-purified synthetic oligo (FF\_oligo347, 50 pmol) with the same sequence as ssDNA(wt) was used as a molecular size marker. The band intensities were measured by Fiji software (51). The intensities were normalized to the intensity of the marker

oligo, and normalized intensities were used to calculate the amount of ssDNA in each sample.

### Induction of cells and plating assays

For each experiment, three transformants were separately inoculated in Luria broth (LB) media plus appropriate antibiotics and grown overnight [37°C, 700 revolutions per minute (RPM)] to obtain seed cultures. Unless otherwise noted, inductions were performed by diluting the seed cultures (1:1000) in 2 ml of prewarmed LB plus appropriate antibiotics with or without inducers followed by 24 hours incubation (30°C, 700 RPM). Aliquots of the samples were then serially diluted, and appropriate dilutions were plated on selective media to determine the number of recombinants and viable cells in each culture. For each sample, the recombinant frequency was reported as the mean of the ratio of recombinants to viable cells for three independent replicates.

In all experiments, the number of viable cells was determined by plating aliquots of cultures on LB-plus-spectinomycin plates. LB-plus-kanamycin plates were used to determine the number of recombinants in the *kanR* reversion assay. For the *galK* reversion assay (Fig. 2), the numbers of *galK*<sub>ON</sub> recombinants were determined by plating the cells on MOPS EZ rich-defined media (Teknova) plus galactose (0.2%). The numbers of *galK*<sub>OFF</sub> recombinants were determined by plating the cells on MOPS EZ rich-defined media plus glycerol (0.2%) plus 2-DOG (2%). For the experiment shown in Fig. 3, the numbers of *kanR*<sub>ON</sub> *galK*<sub>ON</sub> and *kanR*<sub>OFF</sub> *galK*<sub>OFF</sub> cells were determined by using LB-plus-kanamycin plates and MOPS EZ rich-defined media plus glycerol (0.2%), 2-DOG (2%), and D-biotin (0.01%), respectively. The numbers of *kanR*<sub>ON</sub> *galK*<sub>OFF</sub> cells were determined by plating the cells on MOPS EZ-rich defined media plus glycerol (0.2%), 2-DOG (2%), kanamycin, and D-biotin (0.01%).

For the light-inducible SCRIBE experiment (Fig. 4A), induction was performed with white light (using the built-in fluorescent lamp in a VWR 1585 shaker incubator). The “dark” condition was achieved by wrapping aluminum foil around the tubes. Growth of and sampling from these cultures were performed as described earlier.

### LacZ assay

Overnight seed cultures were diluted (1:1000) in prewarmed LB plus appropriate antibiotics and inducers [with different concentrations of aTc or without aTc (Fig. 5, A to C) and with all the four possible combinations of aTc (100 ng/ml) and AHL (50 ng/ml) (Fig. 5, D to F)] and incubated for 24 hours (30°C, 700 RPM). These cultures were then diluted (1:50) in prewarmed LB plus appropriate antibiotics with or without IPTG (1 mM) and incubated for 8 hours (37°C, 700 RPM). To measure LacZ activity, 60  $\mu$ l of each culture was mixed with 60  $\mu$ l of B-PER II reagent (Pierce Biotechnology) and fluorescein di- $\beta$ -D-galactopyranoside (FDG) (0.05 mg/ml final concentration). The fluorescence signal (absorption/emission: 485/515) was monitored in a plate reader with continuous shaking for 2 hours. The LacZ activity was calculated by

normalizing the rate of FDG hydrolysis (obtained from fluorescence signal) to the initial optical density. For each sample, LacZ activity was reported as the mean of three independent biological replicates.

## REFERENCES AND NOTES

- N. Goldman *et al.*, Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **494**, 77–80 (2013). doi: [10.1038/nature11875](#); pmid: [23354052](#)
- G. M. Church, Y. Gao, S. Kosuri, Next-generation digital information storage in DNA. *Science* **337**, 1628 (2012). doi: [10.1126/science.1226355](#); pmid: [22903519](#)
- M. C. Inniss, P. A. Silver, Building synthetic memory. *Curr. Biol.* **23**, R812–R816 (2013). doi: [10.1016/j.cub.2013.06.047](#); pmid: [24028965](#)
- T. S. Gardner, C. R. Cantor, J. J. Collins, Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000). doi: [10.1038/35002131](#); pmid: [10659857](#)
- J. W. Kotula *et al.*, Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4838–4843 (2014). doi: [10.1073/pnas.1321321111](#); pmid: [24639514](#)
- D. R. Burrill, M. C. Inniss, P. M. Boyle, P. A. Silver, Synthetic memory circuits for tracking human cell fate. *Genes Dev.* **26**, 1486–1497 (2012). doi: [10.1101/gad.189035.112](#); pmid: [22751502](#)
- B. P. Kramer *et al.*, An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.* **22**, 867–870 (2004). doi: [10.1038/nbt980](#); pmid: [15184906](#)
- C. M. Ajo-Franklin *et al.*, Rational design of memory in eukaryotic cells. *Genes Dev.* **21**, 2271–2276 (2007). doi: [10.1101/gad.1586107](#); pmid: [17875664](#)
- P. Siuti, J. Yazbek, T. K. Lu, Synthetic circuits integrating logic and memory in living cells. *Nat. Biotechnol.* **31**, 448–452 (2013). doi: [10.1038/nbt.2510](#); pmid: [23396014](#)
- J. Bonnet, P. Subsoontorn, D. Endy, Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8884–8889 (2012). doi: [10.1073/pnas.1202344109](#); pmid: [22615351](#)
- J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, D. Endy, Amplifying genetic logic gates. *Science* **340**, 599–603 (2013). doi: [10.1126/science.1232758](#); pmid: [23539178](#)
- A. E. Friedland *et al.*, Synthetic gene networks that count. *Science* **324**, 1199–1202 (2009). doi: [10.1126/science.1172005](#); pmid: [19478183](#)
- H. M. Ellis, D. Yu, T. DiTizio, D. L. Court, High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6742–6746 (2001). doi: [10.1073/pnas.121164898](#); pmid: [11381128](#)
- N. Costantino, D. L. Court, Enhanced levels of  $\lambda$  Red-mediated recombinants in mismatch repair mutants. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15748–15753 (2003). doi: [10.1073/pnas.2434959100](#); pmid: [14673109](#)
- D. Yu *et al.*, An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5978–5983 (2000). doi: [10.1073/pnas.100127597](#); pmid: [10811905](#)
- J. A. Sawitzke *et al.*, Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. *J. Mol. Biol.* **407**, 45–59 (2011). doi: [10.1016/j.jmb.2011.01.030](#); pmid: [21256136](#)
- B. Swingle *et al.*, Oligonucleotide recombination in Gram-negative bacteria. *Mol. Microbiol.* **75**, 138–148 (2010). doi: [10.1111/j.1365-2958.2009.06976.x](#); pmid: [19943907](#)
- S. Datta, N. Costantino, X. Zhou, D. L. Court, Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1626–1631 (2008). doi: [10.1073/pnas.0709089105](#); pmid: [18230724](#)
- T. Yee, T. Furuchi, S. Inouye, M. Inouye, Multicopy single-stranded DNA isolated from a gram-negative bacterium, *Mycococcus xanthus*. *Cell* **38**, 203–209 (1984). doi: [10.1016/0092-8674\(84\)90541-5](#); pmid: [6088065](#)
- B. C. Lampson, M. Inouye, S. Inouye, Retrons, msDNA, and the bacterial genome. *Cytogenet. Genome Res.* **110**, 491–499 (2005). doi: [10.1159/000084982](#); pmid: [16093702](#)
- D. Lim, W. K. Maas, Reverse transcriptase-dependent synthesis of a covalently linked, branched DNA-RNA compound in *E. coli* B. *Cell* **56**, 891–904 (1989). doi: [10.1016/0092-8674\(89\)90693-4](#); pmid: [2466573](#)
- R. Lutz, H. Bujard, Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I<sub>1</sub>-I<sub>2</sub> regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997). doi: [10.1093/nar/25.6.1203](#); pmid: [9092630](#)
- P. L. Sharma, V. Nurpeisov, R. F. Schinazi, Retrovirus reverse transcriptases containing a modified YXDD motif. *Antivir. Chem. Chemother.* **16**, 169–182 (2005).pmid: [16004080](#)
- J. R. Mao, M. Shimada, S. Inouye, M. Inouye, Gene regulation by antisense DNA produced in vivo. *J. Biol. Chem.* **270**, 19684–19687 (1995). doi: [10.1074/jbc.270.34.19684](#); pmid: [7544343](#)
- S. Warming, N. Costantino, D. L. Court, N. A. Jenkins, N. G. Copeland, Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.* **33**, e36 (2005). doi: [10.1093/nar/gni035](#); pmid: [15731329](#)
- R. Ohlendörf, R. R. Vidavski, A. Eldar, K. Moffat, A. Möglich, From dusk till dawn: One-plasmid systems for light-regulated gene expression. *J. Mol. Biol.* **416**, 534–542 (2012). doi: [10.1016/j.jmb.2012.01.001](#); pmid: [22245580](#)
- M. S. Huen *et al.*, The involvement of replication in single stranded oligonucleotide-mediated gene repair. *Nucleic Acids Res.* **34**, 6183–6194 (2006). doi: [10.1093/nar/gkl852](#); pmid: [17088285](#)
- A. R. Poteete, Involvement of DNA replication in phage lambda Red-mediated homologous recombination. *Mol. Microbiol.* **68**, 66–74 (2008). doi: [10.1111/j.1365-2958.2008.06133.x](#); pmid: [18333884](#)
- A. R. Poteete, Involvement of *Escherichia coli* DNA replication proteins in phage Lambda Red-mediated homologous recombination. *PLOS ONE* **8**, e67440 (2013). doi: [10.1371/journal.pone.0067440](#); pmid: [23840702](#)
- B. E. Dutra, V. A. Suter Jr., S. T. Lovett, RecA-independent recombination is efficient but limited by exonucleases. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 216–221 (2007). doi: [10.1073/pnas.0608293104](#); pmid: [17182742](#)
- S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, R. Weiss, A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130–1134 (2005). doi: [10.1038/nature03461](#); pmid: [15858574](#)
- S. Ausländer, D. Ausländer, M. Müller, M. Wieland, M. Fussenegger, Programmable single-cell mammalian biocomputers. *Nature* **487**, 123–127 (2012).pmid: [22722847](#)
- R. Daniel, J. R. Rubens, R. Sarpeshkar, T. K. Lu, Synthetic analog computation in living cells. *Nature* **497**, 619–623 (2013). doi: [10.1038/nature12148](#); pmid: [23676681](#)
- E. J. Olson, L. A. Hartsough, B. P. Landry, R. Shroff, J. J. Tabor, Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. *Nat. Methods* **11**, 449–455 (2014). doi: [10.1038/nmeth.2884](#); pmid: [24608181](#)
- J. A. Mosberg, C. J. Gregg, M. J. Lajoie, H. H. Wang, G. M. Church, Improving lambda red genome engineering in *Escherichia coli* via rational removal of endogenous nucleases. *PLOS ONE* **7**, e44638 (2012). doi: [10.1371/journal.pone.0044638](#); pmid: [22957093](#)
- S. Miyata, A. Ohshima, S. Inouye, M. Inouye, In vivo production of a stable single-stranded cDNA in *Saccharomyces cerevisiae* by means of a bacterial retron. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5735–5739 (1992). doi: [10.1073/pnas.89.13.5735](#); pmid: [1378616](#)
- O. Mirochnitchenko, S. Inouye, M. Inouye, Production of single-stranded DNA in mammalian cells by means of a bacterial retron. *J. Biol. Chem.* **269**, 2380–2383 (1994).pmid: [7507924](#)
- S. Brakmann, S. Grzeszik, An error-prone T7 RNA polymerase mutant generated by directed evolution. *ChemBioChem* **2**, 212–219 (2001). doi: [10.1002/1439-7633\(20010302\)2:3<212::AID-CBIC212>3.0.CO;2-R](#); pmid: [11828447](#)
- J. D. Roberts, K. Bebenek, T. A. Kunkel, The accuracy of reverse transcriptase from HIV-1. *Science* **242**, 1171–1173 (1988). doi: [10.1126/science.2460925](#); pmid: [2460925](#)
- K. Bebenek, J. Abbotts, S. H. Wilson, T. A. Kunkel, Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots. *J. Biol. Chem.* **268**, 10324–10334 (1993).pmid: [7683675](#)
- K. M. Esvelt, J. C. Carlson, D. R. Liu, A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499–503 (2011). doi: [10.1038/nature09929](#); pmid: [21478873](#)
- Y. Amir *et al.*, Universal computing by DNA origami robots in a living animal. *Nat. Nanotechnol.* **9**, 353–357 (2014). doi: [10.1038/nnano.2014.58](#); pmid: [24705510](#)
- L. Qian, E. Winfree, J. Bruck, Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368–372 (2011). doi: [10.1038/nature10262](#); pmid: [21776082](#)
- G. Seelig, D. Soloveichik, D. Y. Zhang, E. Winfree, Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585–1588 (2006). doi: [10.1126/science.1132493](#); pmid: [17158324](#)
- P. W. Rothmund, Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006). doi: [10.1038/nature04586](#); pmid: [16541064](#)
- S. M. Douglas *et al.*, Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**, 414–418 (2009). doi: [10.1038/nature08016](#); pmid: [19458720](#)
- S. M. Douglas, I. Bachelet, G. M. Church, A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **335**, 831–834 (2012). doi: [10.1126/science.1214081](#); pmid: [22344439](#)
- S. M. Chirieleison, P. B. Allen, Z. B. Simpson, A. D. Ellington, X. Chen, Pattern transformation with DNA circuits. *Nat. Chem.* **5**, 1000–1005 (2013). doi: [10.1038/nchem.1764](#); pmid: [24256862](#)
- C. A. Brosey *et al.*, A new structural framework for integrating replication protein A into DNA processing machinery. *Nucleic Acids Res.* **41**, 2313–2327 (2013). doi: [10.1093/nar/gks1332](#); pmid: [23303776](#)
- S. Kosuri, G. M. Church, Large-scale de novo DNA synthesis: Technologies and applications. *Nat. Methods* **11**, 499–507 (2014). doi: [10.1038/nmeth.2918](#); pmid: [24781323](#)
- J. Schindelin *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi: [10.1038/nmeth.2019](#); pmid: [22743772](#)

## ACKNOWLEDGMENTS

This work was supported by the NIH New Innovator Award (1DP2OD008435), NIH National Centers for Systems Biology (1P50GM098792), the U.S. Office of Naval Research (N000141310424), and the Defense Advanced Research Projects Agency. Sequencing data have been deposited in GenBank with accession numbers KM923743 to KM923754.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/346/6211/1256272/suppl/DC1](http://www.sciencemag.org/content/346/6211/1256272/suppl/DC1)  
Supplementary Text  
Figs. S1 to S6  
Tables S1 to S4  
References (52–61)

19 May 2014; accepted 16 October 2014  
10.1126/science.1256272



## RESEARCH ARTICLE SUMMARY

## DNA RECOMBINATION

# Recombination initiation maps of individual human genomes

Florencia Pratto,\* Kevin Brick,\* Pavel Khil,\* Fatima Smagulova,  
Galina V. Petukhova,† R. Daniel Camerini-Otero†

**INTRODUCTION:** The dramatic events of meiotic recombination culminate in the exchange of genetic information between parental chromosomes and ensure the production of genetically distinct gametes. Recombination is initiated by the formation of programmed DNA double-strand breaks (DSBs), and most DSBs occur at discrete hotspots defined by the DNA binding specificity of the PRDM9 protein. The tandem array of PRDM9 zinc fingers that binds DNA is highly polymorphic, and different variants have different DNA binding preferences. Subsequent to binding, PRDM9 is thought to modify the local chromatin environment and to recruit SPO11 for DSB formation. Meiotic DSBs are predominantly repaired through homologous recombination, giving rise to

either genetic crossovers, in which a reciprocal genetic exchange occurs between homologous chromosomes, or noncrossovers.

**RATIONALE:** Despite recent progress in our understanding of recombination hotspot formation, the initiation of recombination remains poorly understood. Current approaches to study the early steps of meiotic recombination in humans primarily detect genetic crossovers, which is only one of the possible outcomes of DSB repair. Furthermore, these methods are limited by resolution, sex and population averaging, or an inability to extend the analysis genome-wide. To overcome these limitations, we built and analyzed high-resolution, individual-specific maps of meiotic DSBs in the human genome.

**RESULTS:** We report the maps of meiotic DSBs in five males: two homozygous for the most common *PRDM9* allele (*PRDM9<sub>A</sub>*) and three heterozygous for the *PRDM9<sub>A</sub>* allele and for the less frequent *PRDM9<sub>B</sub>* or *PRDM9<sub>C</sub>* alleles. We find that *PRDM9<sub>A</sub>* and *PRDM9<sub>B</sub>* define similar DSB hotspots, whereas the *PRDM9<sub>C</sub>* allele has a distinct specificity. A comparison of DSB hotspot maps with linkage disequilibrium (LD)-based estimates of recombination rates in the human population indicates that the LD map is a superimposition of *PRDM9* allele-specific DSB maps and that the contribution of individual maps is proportional to the *PRDM9* allele frequency in modern Africans. In individuals with identical *PRDM9* alleles, over 5% of DSB hotspots vary in strength, yet less than half of this variation could be explained by

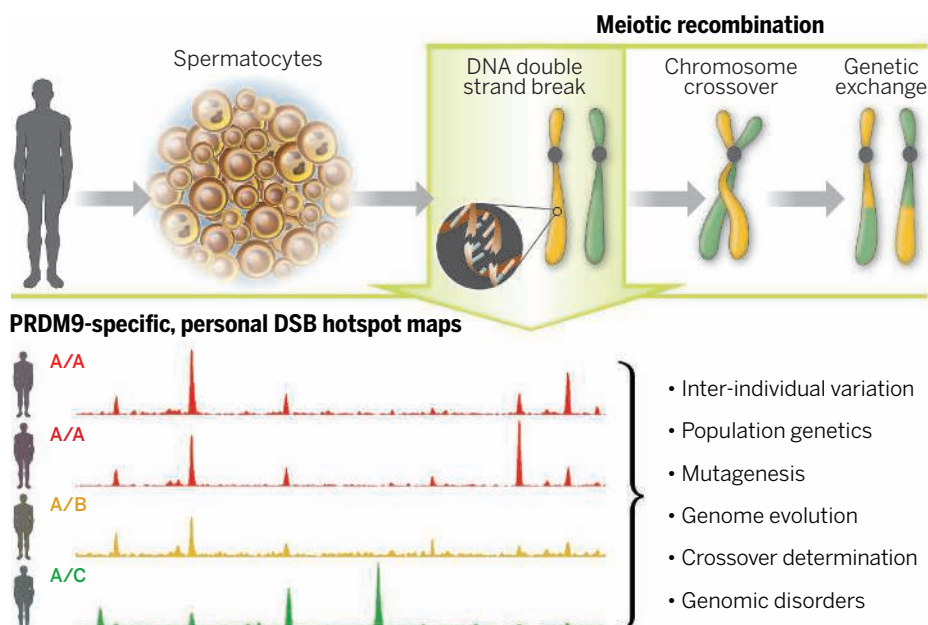
## ON OUR WEB SITE

Read the full article  
at <http://dx.doi.org/10.1126/science.1256442>

sequence variation at putative PRDM9 binding sites. We also find that PRDM9 heterozygosity affects hotspot strength. In human

males, DSBs, like crossovers, occur more frequently at subtelomeric regions, and the crossover rate is directly proportional to our estimate of DSB frequency. This indicates that DSB initiation frequency is a major driver of the crossover rate in human males. We detect distinct signatures of GC-biased gene conversion and of recombination-coupled mutagenesis at DSB hotspots. In addition, DSB hotspots are enriched at the breakpoints of copy number variants that arise via homology-mediated mechanisms. Such variants may give rise to genomic disorders, and indeed, we find that meiotic DSBs defined by *PRDM9<sub>A</sub>* often coincide with disease-associated chromosomal breakpoints.

**CONCLUSIONS:** Our genome-wide recombination initiation maps in individual human males offer unprecedented resolution of recombination initiation sites defined by specific alleles of the PRDM9 protein. Our analysis indicates that DSB frequency is a primary determinant of crossover frequency and that factors other than PRDM9 modulate the frequency of recombination initiation. We also find that meiotic DSB repair and subsequent recombination affect the genome sequence both locally and at the level of structural rearrangements. Taken together, these data provide a foundation for future studies of genetic recombination, meiosis, and genome stability. ■



**Meiotic DSB maps in human male individuals.** In meiotic cells, the formation of programmed meiotic DSBs facilitates the subsequent exchange of genetic material between parental homologous chromosomes. All current methods to study the sites of meiotic recombination rely on detection of these genetic exchanges. The DMC1 protein binds to DNA around meiotic DSBs, and in this work, we used testis biopsies from individual males to pull down the DNA bound by the DMC1 protein. We thus identified the sites of meiotic DSBs in five individual males. Analysis and comparison of the resultant PRDM9-specific, personal genome-wide maps offers insights into the mechanisms that initiate meiosis, genome evolution, crossover formation, human population structure, and predisposition to genomic disorders.

The list of author affiliations is available in the full article online.\*These authors contributed equally to this work.  
†Corresponding author. E-mail: [rdcamerini@mail.nih.gov](mailto:rdcamerini@mail.nih.gov) (R.D.C.-O.); [galina.petukhova@usuh.edu](mailto:galina.petukhova@usuh.edu) (G.V.P.)  
Cite this article as: Pratto et al., *Science* **346**, 1256442 (2014). DOI: 10.1126/science.1256442

## RESEARCH ARTICLE

## DNA RECOMBINATION

# Recombination initiation maps of individual human genomes

Florencia Pratto,<sup>1\*</sup> Kevin Brick,<sup>1\*</sup> Pavel Khil,<sup>1\*</sup> Fatima Smagulova,<sup>2†</sup> Galina V. Petukhova,<sup>2‡</sup> R. Daniel Camerini-Otero<sup>1‡</sup>

DNA double-strand breaks (DSBs) are introduced in meiosis to initiate recombination and generate crossovers, the reciprocal exchanges of genetic material between parental chromosomes. Here, we present high-resolution maps of meiotic DSBs in individual human genomes. Comparing DSB maps between individuals shows that along with DNA binding by PRDM9, additional factors may dictate the efficiency of DSB formation. We find evidence for both GC-biased gene conversion and mutagenesis around meiotic DSB hotspots, while frequent colocalization of DSB hotspots with chromosome rearrangement breakpoints implicates the aberrant repair of meiotic DSBs in genomic disorders. Furthermore, our data indicate that DSB frequency is a major determinant of crossover rate. These maps provide new insights into the regulation of meiotic recombination and the impact of meiotic recombination on genome function.

Genetic variation in humans is shaped by homologous recombination. Meiotic recombination itself is required for correct chromosome segregation in gametes; however, the initiation of recombination in humans remains poorly understood. Recombination events are tightly clustered in 1- to 2-kb-wide hotspots whose position is primarily determined by the histone-lysine *N*-methyltransferase PRDM9 protein (1–6). PRDM9 is one of the most rapidly evolving genes in humans, for which dozens of variants have been described (7–9), and allelic variants have been shown to specify different sets of hotspots (4). The PRDM9 protein binds DNA through a highly polymorphic tandem array of zinc fingers (ZnF) and is then thought to recruit the recombination initiation complex that includes meiotic recombination protein SPO11, the protein that introduces meiotic double-strand breaks (DSBs). These DSBs are subsequently repaired by homologous recombination to give rise to either genetic crossovers, in which a reciprocal genetic exchange occurs between homologous chromosomes, or noncrossovers (10). Designation of a subset of events as crossovers is a tightly regulated process, but the determinants of whether any particular DSB will become a crossover or not remain largely unknown (11).

Several approaches have been used to study recombination hotspots in humans. The most detailed maps of human recombination have been generated by means of computational inference

of recombination rates from patterns of linkage disequilibrium (LD) in the human population (12–15). These maps do not, however, provide information about recombination rates in individuals. Single-nucleotide polymorphism (SNP) genotyping in human pedigrees has been used to identify crossovers, but the precision of crossover mapping is tens of kilobases (16–20). Until recently, the study of meiotic recombination hotspots in individuals required sperm genotyping (21), a method that can define crossover sites with nucleotide resolution but cannot be used for genome-wide analyses. Single-cell sequencing techniques have facilitated the construction of genome-wide crossover maps from individual sperm and oocytes (22–25); however, such approaches currently lack the resolution to perform fine-scale analysis and hotspot detection. Furthermore, these approaches rely on the identification of crossovers, which are just one of the possible outcomes of recombination. In mammals, only ~10% of DSBs are repaired as crossovers (26); therefore, the vast majority of recombination events in meiosis remain unexplored.

We overcome these limitations by generating genome-wide maps of meiotic recombination initiation sites in individual human males. Our approach combines hotspot resolution that is comparable with sperm genotyping, with the gender, individual, and PRDM9-allele specificity that LD-based methods lack. In addition, we detect all sites where recombination can occur, independent of the subsequent repair pathway. Although PRDM9 dictates hotspot locations, we find that ~5% of hotspots are polymorphic between individuals with identical PRDM9 alleles. Sequence polymorphism at PRDM9 binding sites explains less than half of this variation. Through analysis of the DNA polymorphism spectrum at hotspots, we identified distinct signatures of GC-biased gene conversion and of recombination-mediated mutagenesis. We

find evidence for a role of ectopic recombination in gross chromosomal rearrangements and identify 726 new potential rearrangement breakpoints. Last, this first analysis of the recombination initiation landscape establishes that like crossovers, DSBs occur more frequently at subtelomeric regions. This suggests that initiation frequency is a major driver of crossover rate in human males.

## Genome-wide DSB hotspot map in humans

To create a representative overview of the recombination initiation landscape in human males, we generated high-resolution maps of meiotic DSB hotspots from four unrelated individuals (Fig. 1, fig. S1, and table S1). We performed chromatin immunoprecipitation followed by single-stranded DNA sequencing (SSDS) (27) to identify DNA fragments associated with DMC1 (meiotic recombination protein DMC1/LIM15 homolog), a specific marker of meiotic DSBs. The number of DMC1-associated single-stranded DNA (ssDNA) fragments provides an estimate of DSB frequency, although this estimate could be affected by the relative lifetime of ssDNA intermediates and by differences in DMC1 loading at individual hotspots. Two men in our sample set were homozygous for the most common PRDM9 allele (PRDM9<sub>A</sub>; individuals AA<sub>1</sub> and AA<sub>2</sub>), and three were heterozygous for the PRDM9<sub>A</sub> allele and for the less frequent variants PRDM9<sub>B</sub> (AB<sub>1</sub>, AB<sub>2</sub>) and PRDM9<sub>C</sub> (AC). In total, we identified up to 38,946 DSB hotspots per individual (Fig. 1B and table S1). This number is substantially higher than the 15,000 to 20,000 hotspots identified in mouse (4), which is perhaps a reflection of the twofold higher recombination rate in humans (28). The SSDS signal was three- to sevenfold higher on the sex chromosomes than on the autosomes (fig. S1), and this may reflect continuous DSB formation (29) or extended DSB life span on the sex chromosomes (30, 31). Autosomal DSB hotspot strength varied by over three orders of magnitude (fig. S2), and ~100 hotspots (fig. S2, table S3, and data file S1) were stronger than the recombination hotspot with the highest known crossover rate (9, 32).

The PRDM9 protein defines most hotspot sites in mice (4). Consistent with this role, 89% of human DSB hotspots were found at the same locations in the two AA individuals (fig. S3). A similar proportion of hotspots in the AB individual overlapped AA<sub>1</sub>/AA<sub>2</sub> hotspots (88%), suggesting that the PRDM9<sub>A</sub>-like PRDM9<sub>B</sub> allele does not specify a distinct set of hotspots (33) (fig. S3). Only 43% of DSB hotspots in the AC individual overlap PRDM9<sub>A</sub>-defined hotspots (Fig. 1B); therefore, the remaining 56% (19,330) are likely PRDM9<sub>C</sub>-defined. At the local level, many properties of hotspots, regardless of the PRDM9 allele, remain conserved between human and mouse (figs. S4 and S5) (33). Common hotspot features include a purine-pyrimidine skew around hotspot centers and a local increase in GC content. Furthermore, as in mouse and consistent with the role of PRDM9 as a histone H3K4 trimethyltransferase (34), most human DSB hotspots coincide with H3K4me3 in testis (57%) (fig. S6).

<sup>1</sup>National Institute of Diabetes, Digestive and Kidney Diseases, NIH, Bethesda, MD, USA. <sup>2</sup>Department of Biochemistry and Molecular Biology, Uniformed Services University of Health Sciences, Bethesda, MD, USA.

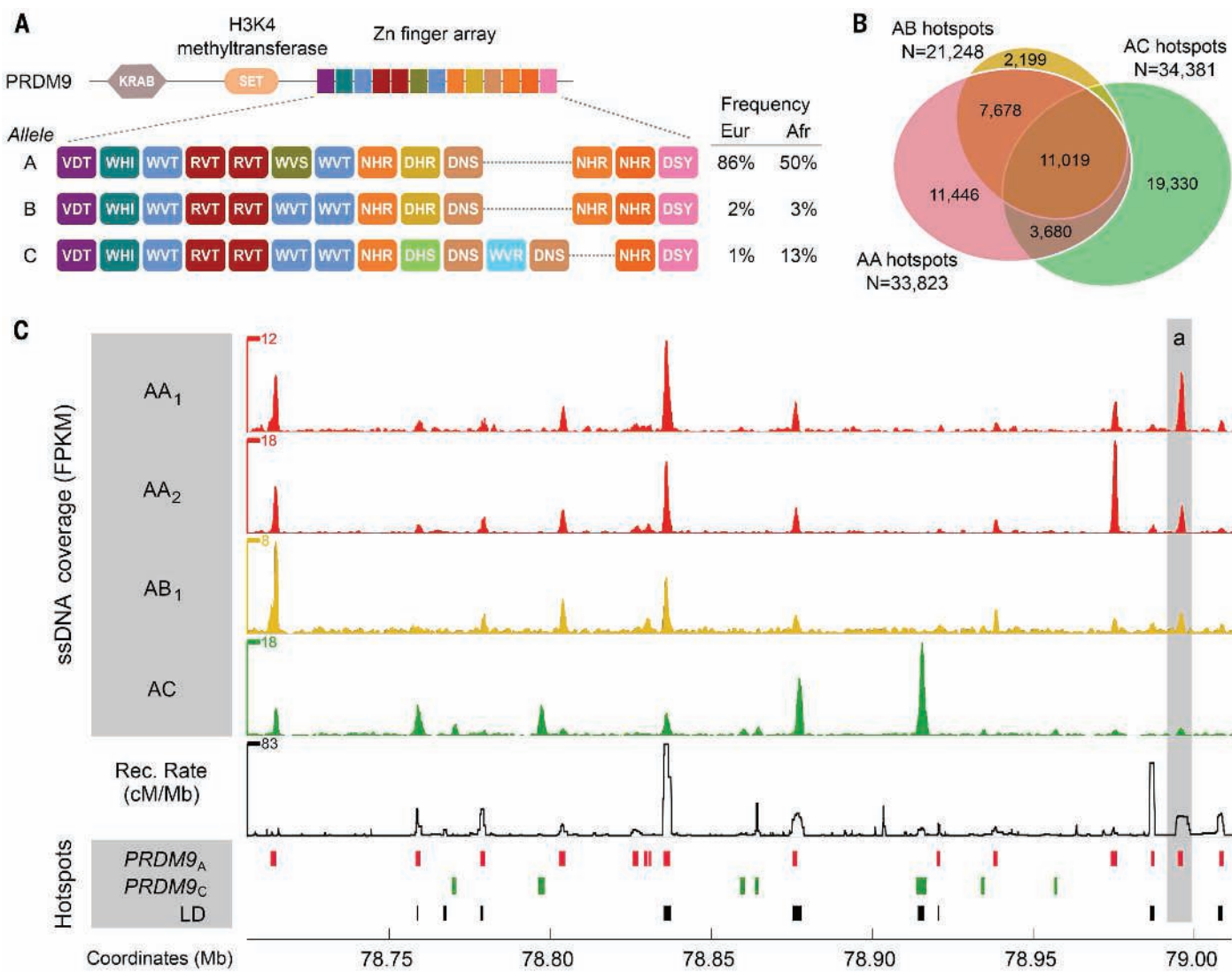
\*These authors contributed equally to this work. †Present address: Inserm UMR1085-Irset, 263 Ave du Général Leclerc, 35042, Rennes, France. ‡Corresponding author. E-mail: rdcamerini@mail.nih.gov (R.D.C.-O.); galina.petukhova@usuhs.edu (G.V.P.)



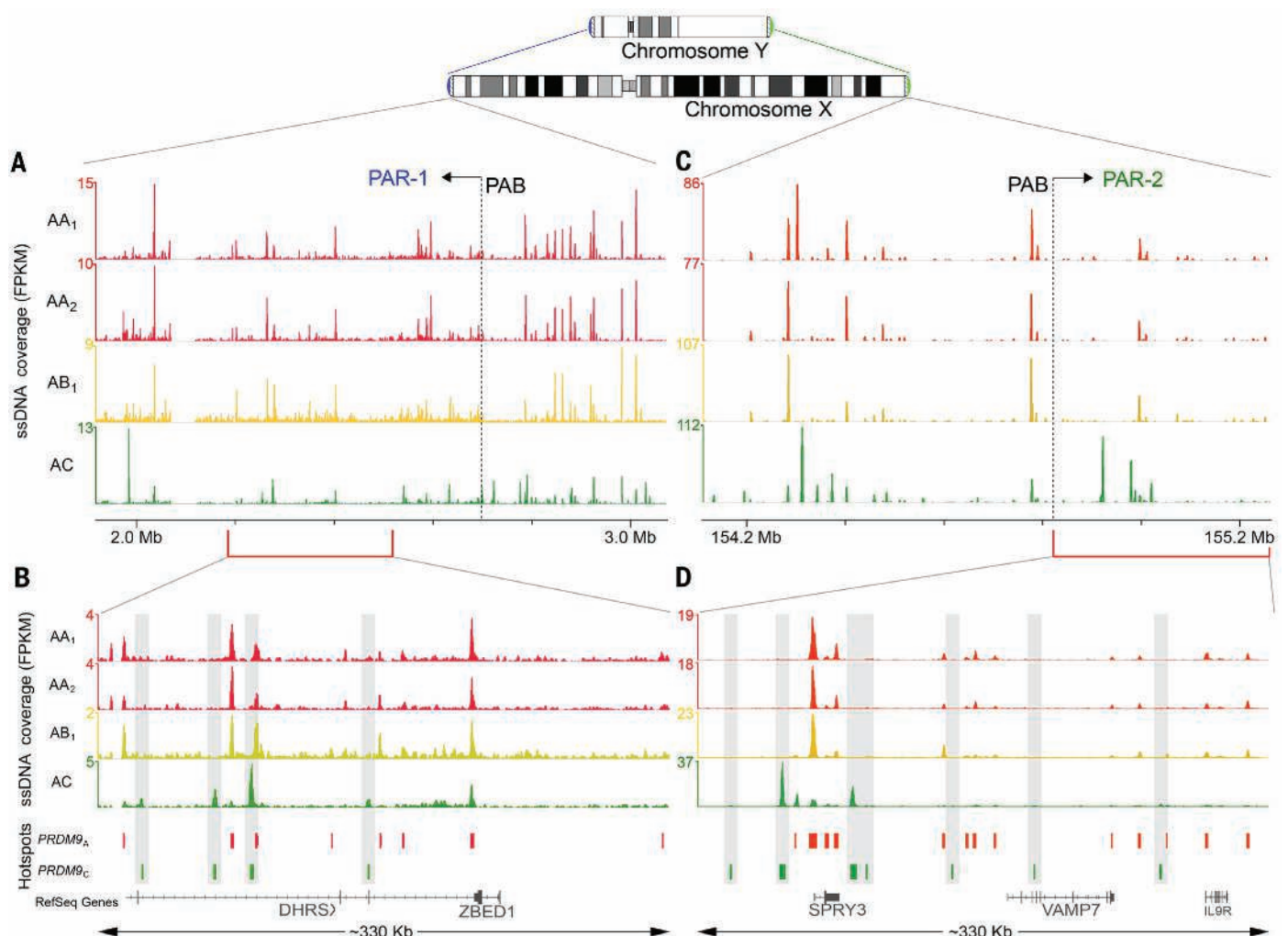
Consistent with the different DNA binding specificities of the *PRDM9*<sub>A</sub> and the *PRDM9*<sub>C</sub> alleles, we identified distinct consensus motifs enriched at the centers of *PRDM9*<sub>A</sub>- and *PRDM9*<sub>C</sub>-defined hotspots (fig. S7). Each motif matches the predicted *PRDM9* binding site for the respective *PRDM9* allele but not that of the other allele (fig. S7). Furthermore, the *PRDM9*<sub>A</sub> motif, but not the *PRDM9*<sub>C</sub> motif, is highly similar to a 13-oligomer motif previously found to be enriched at recombination hotspots (LD-hotspots) (fig. S7) (13, 35). This is consistent with the widespread predominance of the *PRDM9*<sub>A</sub> allele in human populations [84% in European, 50% in African populations (9)]. In turn, the *PRDM9*<sub>C</sub> motif is highly similar to a 17-oligomer motif, found at LD-defined recombination hotspots used in the African American but not in the European population (20, 35). This agrees

with the increased prevalence of the *PRDM9*<sub>C</sub> allele in Africans (13%) as compared with Europeans (~1%) (9). To ensure proper segregation of the sex chromosomes during male meiosis, a crossover must form in the short regions of sequence homology shared between the X and Y chromosomes [pseudoautosomal regions (PARs)]. In mouse, there is a broad and extremely intense DSB signal adjacent to the pseudo-autosomal boundary (PAB) Furthermore, this signal and several hotspots outside of the PAR but immediately adjacent to it appear to be formed independently of *PRDM9* (4). Unlike in mouse, we did not observe a prominent DSB cluster near either human PAB (Fig. 2, A and B), and *PRDM9* allele-specific hotspot formation was observed in both PAR regions (Fig. 2, C and D). Thus, it is unlikely that *PRDM9*-independent

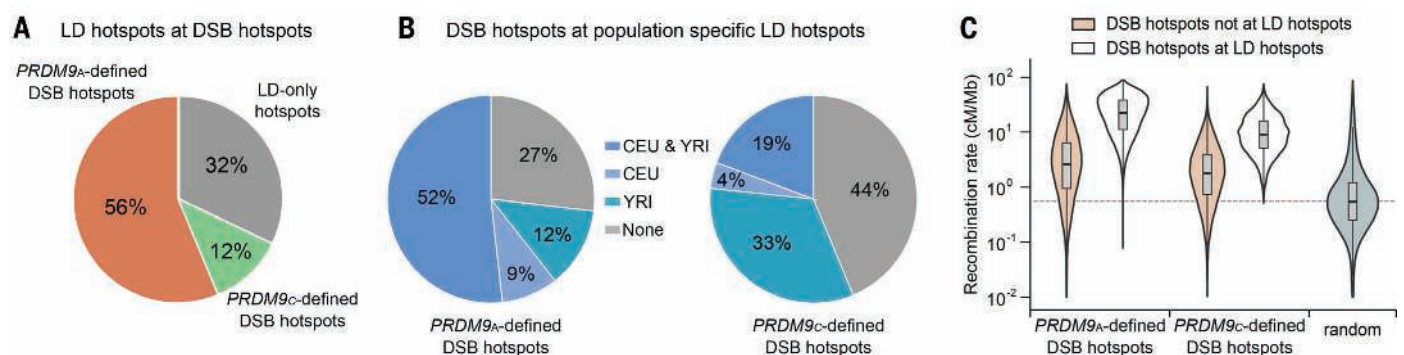
DSB formation near the PAB is a conserved mechanism to ensure a mandatory crossover in the PAR. It is possible, however, that *PRDM9*-independent DSB clusters are located in the very distal part of the human PARs close to the telomere. These regions are poorly assembled and replete with repetitive DNA and therefore cannot be analyzed by using high-throughput sequencing. **Contribution of individual *PRDM9* alleles to the LD-based recombination rate map** LD-based methods have provided the most comprehensive maps of human recombination to date (12–15). Because these maps are intrinsically sex- and population-averaged, they cannot distinguish between hotspots defined by different *PRDM9* alleles. In contrast, our approach directly measures DSB frequency in a single male with a



**Fig. 1. Genome-wide distribution of DSB hotspots in human individuals.** (A) ZnF array structure of *PRDM9* alleles in individuals from this study. Population allele frequencies are taken from (9). (B) The overlap between autosomal hotspots found in different individuals. AA hotspots are the hotspots found in either AA<sub>1</sub> or AA<sub>2</sub> individuals. (C) DSB hotspots from AA<sub>1</sub>, AA<sub>2</sub>, AB<sub>1</sub>, and AC individuals in a 300-kb region on chromosome 17. (Top four) Normalized ssDNA coverage in fragments per kilobase per million (FPKM), smoothed by using a sliding window (window, 1 kb; step, 0.1 kb). The baseline of the y axes is 0 FPKM. (Second from bottom) The recombination rate calculated from the population-averaged HapMap data (15). (Bottom) *PRDM9*<sub>A</sub>- and *PRDM9*<sub>C</sub>-defined DSB hotspots and LD-defined hotspots.



**Fig. 2. PRDM9-defined hotspots are found in the human PARs.** (A and B) Overview of ~1-Mb regions on chromosome X overlapping the (A) PAR1 and (B) PAR2 pseudoautosomal boundaries (PABs). Normalized ssDNA coverage is shown as FPKM, smoothed by using a sliding window (window, 1 kb; step, 0.1 kb). The baseline of the y axes is 0 FPKM. Coverage is shown for the AA<sub>1</sub>, AA<sub>2</sub>, AB, and AC individuals. Red and green bars depict PRDM9<sub>A</sub>- and PRDM9<sub>C</sub>-defined hotspots, respectively. (C and D) Closeup of a (C) 330 kb region of the PAR1 and (D) the entire ~330 kb PAR2 illustrating the presence of PRDM9<sub>C</sub>-defined hotspots (highlighted by gray bars) in both PARs.



**Fig. 3. Comparison of LD-based recombination maps and DSB hotspots.** (A) The proportion of LD-hotspots that overlap PRDM9<sub>A</sub>- or PRDM9<sub>C</sub>-defined DSB hotspots. Most LD-hotspots are detected in the DSB hotspot maps. (B) The percentage of PRDM9<sub>A</sub>- (left) and PRDM9<sub>C</sub>-defined (right) hotspots overlapping LD-hotspots in the CEU (European) and YRI (African) populations. Most DSB hotspots are found in the LD maps. (C) Violin plots showing the population-averaged recombination rate at DSB hotspots that overlap or do not overlap LD-hotspots defined in the CEU and YRI populations. 85% of PRDM9<sub>A</sub>-defined and 82% of PRDM9<sub>C</sub>-defined hotspots that do not overlap an LD-hotspot have a recombination rate above random expectation.



known combination of *PRDM9* alleles. We thus explored the contribution of individual *PRDM9* alleles to the LD map.

Overall, we found a good agreement between our individual-specific DSB maps and the LD map (Fig. 3A), with 68% of LD-hotspots coinciding with a DSB hotspot. Consistent with the high frequency of the *PRDM9*<sub>A</sub> allele across human populations, 56% of LD-hotspots coincided with *PRDM9*<sub>A</sub>-defined hotspots (Fig. 3A) (9). The less frequent *PRDM9*<sub>C</sub> allele has also left a considerable footprint on patterns of linkage disequilibrium because 12% of LD-hotspots overlapped *PRDM9*<sub>C</sub>-defined hotspots. These proportions are very close to the frequencies of these *PRDM9* alleles in modern Africans (Fig. 1A) (9), suggesting that the *PRDM9* allelic distribution in modern Africans is similar to that in ancestral human populations. Nevertheless, 32% of LD-hotspots do not overlap a DSB hotspot in any of our maps (Fig. 3A). These “LD-only” hotspots are likely the products of minor *PRDM9* alleles, hotspots that vary between individuals and hotspots that are used more frequently in females than in males.

We next asked what proportion of DSB hotspots can be seen in LD data. Overall, 51.1% of DSB hotspots overlap population-averaged LD-hotspots (15). Because there are substantial differences in *PRDM9* allele frequencies between human populations, we used LD-hotspots inferred from population-specific subsets of HapMap II data (19) [Utah residents with ancestry from northern and western Europe (CEU); Yoruba in Ibadan, Nigeria (YRI)] to assess the DSB overlap. In agreement with the higher prevalence of *PRDM9*<sub>C</sub> in African as compared with European populations, the *PRDM9*<sub>C</sub>-defined DSB hotspots are better represented at YRI-specific hotspots (33% overlap) (Fig. 3B) than at CEU-specific hotspots (4% overlap) (Fig. 3B). Furthermore, at *PRDM9*<sub>C</sub>-defined DSB hotspots, the mean YRI-derived recombination rate is higher than the mean CEU-derived recombination rate (fig. S8). In contrast, *PRDM9*<sub>A</sub>-defined DSB hotspots are well represented in both CEU- and YRI-specific LD-hotspots, with 52% overlapping hotspots in both populations (Fig. 3B). Although the majority of DSB hotspots for each *PRDM9* allele are found at LD-hotspots, 27% of *PRDM9*<sub>A</sub>-defined and 44% of *PRDM9*<sub>C</sub>-defined DSB hotspots are not found at a LD-hotspot in either population (Fig. 3B). We did find, however, that >80% of these DSB hotspots were located in a region with an elevated recombination rate (Fig. 1C, “a,” and Fig. 3C), suggesting that these hotspots have been active in human populations but were simply below the detection thresholds used for LD-hotspot detection. Together, the DSB maps for different *PRDM9* alleles clearly show that the population-averaged LD map is a combination of allele-specific maps.

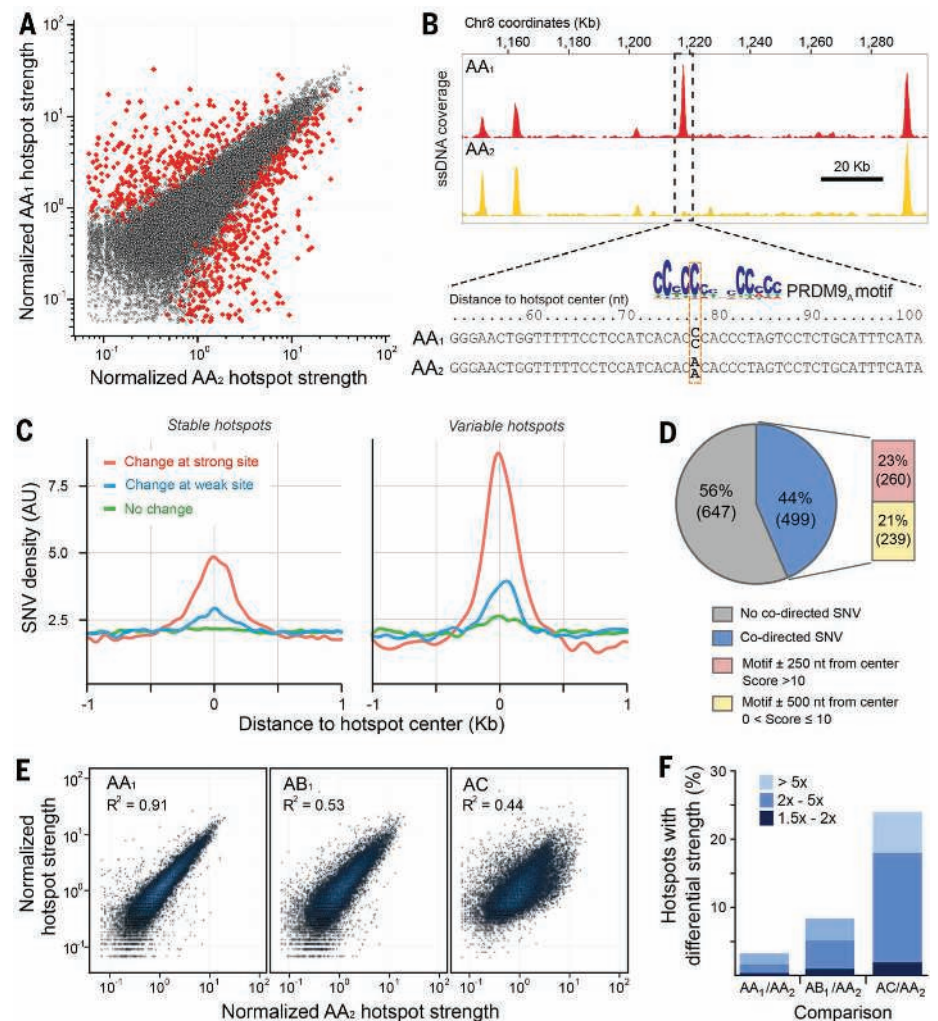
### Interindividual variability in hotspot strength

Although hotspot strength has been shown to vary between individuals sharing the same *PRDM9* genotype (8, 9), these analyses were based on comparisons of crossover frequency at a limited

number of strong hotspots, and the extent of this variation genome-wide is not known. To explore variation in hotspot strength on a genome-wide scale, we compared our DSB maps between the AA<sub>1</sub> and AA<sub>2</sub> individuals (Fig. 4A) (33). Conservatively, we estimated that at least 3.2% (1146) of hotspots varied in strength (1.25- to >30-fold) (fig. S9), although the true proportion may be higher (fig. S9A).

Multiple factors may affect the DSB initiation frequency at a given hotspot. These include the

DNA binding affinity of *PRDM9*, accessibility of the *PRDM9* binding site, and modulation of SPO11 recruitment and DNA cleavage. To determine how interindividual sequence variation at *PRDM9* binding sites affects DSB hotspot strength (Fig. 4B), we performed whole-genome sequencing of the AA<sub>1</sub>, AA<sub>2</sub>, and AB individuals and identified 6.3 million sequence variants that differed from the reference genome (table S4). We then calculated the *PRDM9* DNA-binding motif match scores associated with each variant. Motif-changing



**Fig. 4. Variation in hotspot strength between the AA<sub>1</sub> and AA<sub>2</sub> individuals.** (A) A comparison of *PRDM9*<sub>A</sub>-defined hotspot strength in the AA<sub>1</sub> and AA<sub>2</sub> individuals. Variable hotspots that satisfy our statistical criteria are shown in red. (B) Example of a variable hotspot between the two AA individuals in which a homozygous SNV disrupts a *PRDM9* motif near the hotspot center. The *PRDM9* motif match is better in AA<sub>1</sub> than in AA<sub>2</sub>, and the change in hotspot strength is codirected with the change in motif score. (C) Motif-affecting SNVs are enriched around hotspot centers. The SNV density was smoothed by using a Gaussian kernel estimation and is shown for variants that change the motif score at a strong putative *PRDM9* binding site (red, score > 10), at a weak putative *PRDM9* binding site (blue, 0 < score ≤ 10), or that do not change a putative binding site (green). Only changes >1 bit are considered. The y axis is expressed in arbitrary units (AU). Separate plots for stable (left, n = 35,228 hotspots) and for variable (right, n = 1146 hotspots) hotspots from the AA<sub>1</sub>/AA<sub>2</sub> comparison are shown. (D) The proportion of variable hotspots that contain a codirected motif-affecting SNV near the hotspot centers. (E) *PRDM9*<sub>A</sub>-defined hotspot strength in the AA<sub>1</sub>, AB<sub>1</sub>, and AC individuals compared with the AA<sub>2</sub> individual. We used 15,051 *PRDM9*<sub>A</sub>-defined hotspots common to all four individuals. (F) Quantification of variable *PRDM9*<sub>A</sub>-defined hotspots between individuals. Bar colors indicate fold change.

sequence variants were strongly enriched around the center of variable hotspots (Fig. 4C), and changes in motif score were positively correlated with changes in hotspot strength genome-wide (fig. S10). Furthermore, the enrichment of motif-score changing single-nucleotide variants (SNVs) around hotspot centers is primarily driven by “codirected” SNVs, in which the change in motif score matches the direction of change in hotspot strength (fig. S11).

To estimate the proportion of variable hotspots likely explained by sequence variation at PRDM9 binding sites, we first evaluated the spatial distribution and strength of putative functional PRDM9 sites (sites found inside hotspots). We estimate that >70% of functional PRDM9 binding sites lie within 250 nucleotides (nt) of hotspot centers, while >92% lie within 500 nt (fig. S12A). 99% of hotspots contain a PRDM9 motif match with a score >10 within 250 nt of center (fig. S12, B and C). However, analysis of PRDM9 motif loss through evolution suggests that even motifs with a score between 0 and 5 have been functionally active (fig. S12D). In total, 88% of variable hotspots differ in at least one sequence

position between the genomes of AA<sub>1</sub> and AA<sub>2</sub> (table S5), providing an upper limit of proportion of variable hotspots caused by variation at PRDM9 binding sites. More realistically, between 23% (score > 10 within 250 nt from center) and 44% (score > 0 within 500 nt from center) of variable hotspots are likely explained by sequence differences at PRDM9 binding sites (Fig. 4D). Because sequence variation explains less than 44% of variable hotspots, other factors such as binding site accessibility must strongly affect DSB initiation frequency. The relatively minor impact of sequence variation at DNA binding sites is not unique to meiotic DSB hotspots; for transcription factors, sequence changes at putative DNA binding sites only explain a small fraction of differential transcription factor occupancy (36, 37).

Next, we asked whether variable hotspots detected in the AA<sub>1</sub>/AA<sub>2</sub> comparison were specific to these individuals or rather were polymorphic among humans. Individual-specific hotspots will not have left a footprint on historical recombination rates, and therefore, we compared the LD-defined recombination rate between variable and stable hotspots. We observed little difference

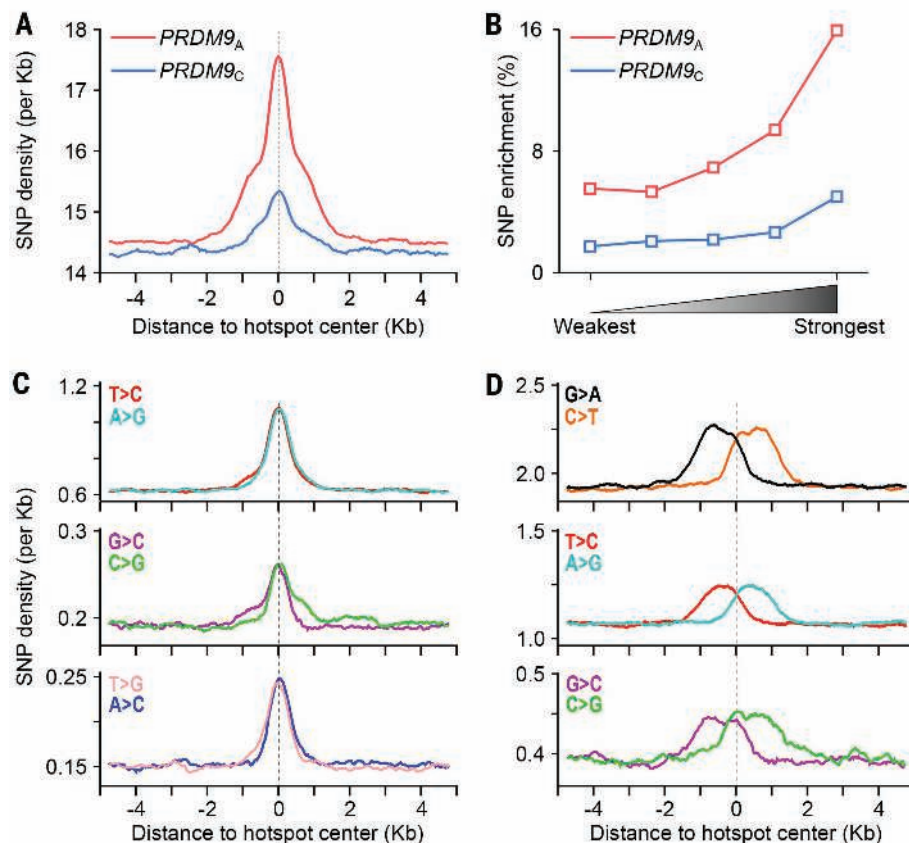
between the mean recombination rates (fig. S13), suggesting that few variable hotspots are specific to AA<sub>1</sub> or AA<sub>2</sub>. In addition, over 60% of codirected SNVs within 250 nt from center are commonly found (minor allele frequency of >0.05) in both the CEU and YRI populations (fig. S14), suggesting that most variable hotspots driven by variation at the PRDM9 binding site are polymorphic in humans.

### PRDM9 heterozygosity modulates hotspot strength

DSB hotspots can vary in strength between individuals homozygous for the PRDM9<sub>A</sub> allele. We next asked whether a heterozygous combination of PRDM9 alleles contributes to hotspot strength variation at PRDM9<sub>A</sub>-defined hotspots. We compared hotspot strength in the AA<sub>1</sub>, AB, and AC individuals to the strength in AA<sub>2</sub> and found a much higher proportion of variable PRDM9<sub>A</sub>-defined hotspots in the AB and AC individuals (8.5 and 24.0%, respectively) compared to AA<sub>1</sub>/AA<sub>2</sub> (3.5%) (Fig. 4, E and F, and fig. S15). This difference is unlikely to be caused by sequence variation in PRDM9 binding sites because the number of potential motif-disrupting sequence variants is similar in the AA and AB individuals (fig. S16 and table S5). In the case of the PRDM9<sub>A</sub>/PRDM9<sub>B</sub> heterozygous individual, it is possible that the PRDM9<sub>B</sub> ZNF array has slightly different binding preferences compared with PRDM9<sub>A</sub>, and therefore, the apparent hotspot strength variation might be just a reflection of these changes in binding affinity. However, we do not observe any substantial changes in PRDM9<sub>B</sub> binding specificity (fig. S3) (33), and such a model cannot account for increased hotspot variation in the PRDM9<sub>A</sub>/PRDM9<sub>C</sub> individual.

One way to explain the increased hotspot variability in the AC individual is interference between neighboring hotspots. Indeed, correlated changes in crossover frequency at neighboring hotspots have been observed in humans (38). In *Saccharomyces cerevisiae*, there is also abundant evidence that the activation of a nearby hotspot can affect DSB frequency at other hotspots in its vicinity (39–42). Chromatin modifications following DSB formation could provide a mechanistic basis for such effects. For example, H2AX phosphorylation occurs rapidly after DSB formation and can span megabase-sized domains in mammals (43). This distance is clearly sufficient to affect nearby hotspots.

Alternatively, interactions between PRDM9 monomers, either direct or mediated by cofactors, may affect the DNA binding activity of PRDM9 and thus explain the increased hotspot strength variance in individuals heterozygous for PRDM9. Such cooperative interactions could modulate hotspot strength without changing binding specificity and may result in partial dominance (44). In the case of PRDM9 binding, we have clear partial dominance of one allele over the other (fig. S17) (4). Although we cannot clearly establish the mechanism of hotspot strength modulation, this effect is not restricted to humans. In a mouse hybrid derived from a 9R13R F1 cross,



**Fig. 5. Signatures of increased genetic diversity at DSB hotspots.** (A) A local increase in SNP density is observed at both PRDM9<sub>A</sub>- (red) and PRDM9<sub>C</sub>-defined (blue) hotspots. All SNPs from the 1000 genomes project were used. (B) The magnitude of SNP enrichment at hotspots is positively correlated with hotspot strength. SNP enrichment is calculated as the SNP density in the central  $\pm 1.5$  kb relative to the mean SNP density in the region from 4 to 5 kb from the hotspot center. (C) Common AT>GC and GC>CG variants are enriched in the  $\pm 0.5$ -kb region around the hotspot center. (D) Rare variants are enriched in a region  $\pm 1.5$  kb around the hotspot center. Only variants enriched at hotspots are shown. Variants exhibit rotational symmetry around the hotspot center.



we also observe increased variance in hotspot strength in heterozygous F1 animals relative to *PRDM9* homozygotes (fig. S18). Taken together, these observations suggest that the presence of a second *PRDM9* allele can influence DSB hotspot strength.

### The mutagenic effect of meiotic recombination

Meiotic recombination influences genome evolution through the shuffling of parental alleles, and broad-scale recombination rates are positively correlated with genetic diversity (45). At finer scales, the recombination rate has also been found to positively correlate with genetic diversity in humans (46, 47). However, the use of polymorphisms themselves to infer LD-defined recombination rates may confound such analyses (15). In order to better understand how meiotic recombination influences genome diversity at a local scale, we explored the patterns of DNA variation around DSB hotspots.

Our initial analyses revealed a local increase in SNP density in the ~3-kb region around both *PRDM9*<sub>A</sub>- and *PRDM9*<sub>C</sub>-defined hotspot centers (Fig. 5A). This local increase in SNP density is likely a direct consequence of meiotic DSBs because the magnitude of enrichment was positively correlated with hotspot strength (Fig. 5B). In addition, SNP enrichment reflects historical hotspot usage, as enrichment at *PRDM9*<sub>A</sub>-defined hotspots was three times greater than at *PRDM9*<sub>C</sub>-defined hotspots. To account for the effects of selection and population history, we investigated the distribution of SNPs with different derived allele frequencies (DAFs) around DSB hotspots. In general, common SNPs represent relatively old mutations that have become established in the population as a result of selection or genetic drift. Rare variants are less likely to be influenced by selection; therefore, they will more accurately reflect the spectrum of mutagenesis events (48). Parsing by DAF revealed two distinct spatial profiles of SNP enrichment, each of which was correlated with hotspot strength (figs. S19 and S20); a signal in the central ±0.5 kb derived primarily from common variants, and a broader signal extending to the ±1.5-kb shoulders of hotspots derived from rare variants.

Among common variants (DAF >0.01), AT>GC, AT>CG, and GC>CG variants were enriched in the narrow central ±0.5-kb region of hotspots (Fig. 5C and fig. S21). This polymorphism spectrum is indicative of GC-biased gene conversion (gBGC) (49), and the 1-kb extent of this signature closely approximates recent measurements of gBGC at mouse hotspots (50). Consistent with gBGC, we observe fixation of and enrichment for GC nucleotides at the hotspot center (fig. S22). The polymorphism spectrum of enriched rare variants (DAF <0.01) was quite different from that observed for common variants. Among rare variants C>T (G>A) transitions, T>C (A>G) transitions, and to a lesser extent C>G (G>C) transversions were enriched in the broad ±1.5-kb region around the hotspot centers (Fig. 5D and fig. S21). Like DNA resection, the rare polymorphism spectrum

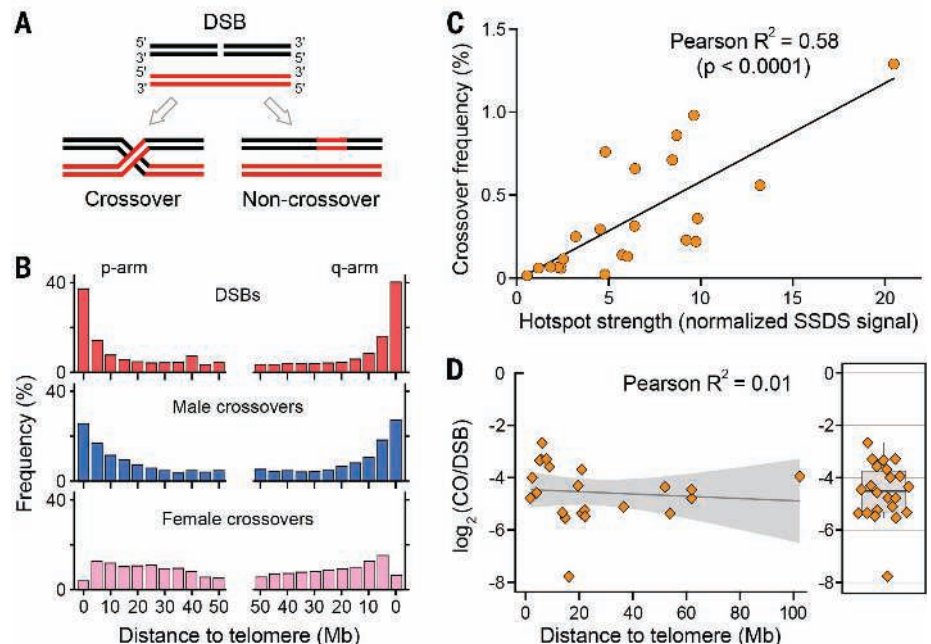
exhibited 180° rotational symmetry around the hotspot center (Fig. 5D and figs. S21 and S23). Together with the strength-dependence of enrichment (fig. S19), this symmetry suggests that these variants arose directly from DSB repair processes. A more detailed analysis of the trinucleotide context of SNPs shows that a majority of rare C>T and G>A variants occurred at ancestral CpG dinucleotides (fig. S24, A and B). Nevertheless, the polarity of these variants around DSBs makes it unlikely that cytosine deamination, a major mutagenic mechanism affecting methylated CpG dinucleotides (51), was the mechanism of their formation (fig. S24C). Exactly which mechanism drives this diversity remains unclear; however, error-prone DNA synthesis by trans-lesion polymerases may have a role in meiotic DSB repair (52). Together, our data show unequivocally that meiotic DSB repair processes increase local genetic diversity through both gBGC and mutagenesis.

### DSB frequency is a major determinant of crossover rate

In meiosis, a DSB can be repaired as either a crossover or as a noncrossover (Fig. 6A) (26). Because the proportion of DSBs resolved as crossovers might vary from hotspot to hotspot (11), the frequency of crossing over need not necessarily reflect the DSB formation rate. We thus asked whether the crossover landscape is largely shaped by variation in crossover/noncrossover

resolution or whether it is mostly determined by the DSB frequency.

It has been well established that crossovers in human males are enriched in subtelomeric regions (Fig. 6B and fig. S25) (16, 18, 19, 22, 23, 25, 53). We found that DSB hotspots were also stronger and more densely spaced in the distal parts of chromosomes (Fig. 6B and figs. S1 and S26). This enrichment spans ~10 Mb, independent of chromosome size (fig. S27). Quantitatively, at the megabase scale the SSDS signal was strongly correlated to the male (Pearson  $R^2 = 0.90$ ,  $n = 22$  intervals of 5 Mb averaged across all chromosomes,  $P < 0.0001$ ) but not to the female crossover frequency (18) ( $R^2 = -0.07$ ,  $n = 22$  intervals,  $P = 0.36$ ) (Fig. 6B). Furthermore, at *PRDM9*<sub>A</sub>-defined hotspots we also found a positive correlation between the SSDS signal and the LD-defined recombination rates (Spearman  $R^2 = 0.2$ ) (33), despite the influence of female recombination and hotspot erosion on these rate estimates. The relationship between crossover frequency and SSDS signal remains strong at the level of individual hotspots. We found that the mean SSDS signal in AA individuals is strongly correlated with the mean crossover frequency determined with sperm genotyping (Pearson  $R^2 = 0.58$ ;  $P < 0.0001$ ) (Fig. 6C and table S3) (9, 32), although the CO:DSB ratio varies among individual hotspots. The preferential resolution of DSBs as COs near telomeres may also contribute to the observed CO distribution, and



**Fig. 6. DSB frequency is correlated with the crossover rate.** (A) Meiotic DSBs can be repaired as either a crossover or a noncrossover. (B) The distribution of *PRDM9*<sub>A</sub>-defined DSBs from the AA<sub>2</sub> (red) individual and broad-scale distribution of (blue) male- and (pink) female-derived crossovers from (18). (C) The average DSB frequency from the two AA individuals is correlated with crossover frequency at 22 *PRDM9*<sub>A</sub>-defined hotspots ( $P < 10^{-4}$ , two tailed  $t$  test). Crossover frequency is taken from (9) and (32). (D) The CO:DSB ratio is not strongly dependent on the distance to the telomere. The CO:DSB ratio for each of the hotspots analyzed in (C) is plotted against the absolute distance of the hotspot to the closest telomere. The shaded region represents the SE for the linear fit line. (Right) Boxplot illustrating that the CO:DSB ratio is highly variable among individual hotspots.

we therefore analyzed the relationship between CO:DSB ratio and proximity to telomeres. On the basis of the limited set of available hotspots, we found no strong evidence that the CO:DSB ratio depends on the distance to the telomere (Fig. 6D), although nonlinear fitting suggests that the CO:DSB ratio may be higher near the telomere (fig. S28A). This increase could be driven by stronger CO hotspots (fig. S28B), which are themselves known to be enriched near telomeres; therefore, more data will be required to establish whether the CO:NCO ratio is elevated in telomere-adjacent regions. Taken together, although biased CO/NCO resolution may contribute, our per-hotspot and genome-wide observations collectively indicate that crossover frequency is largely determined by the rate of DSB formation.

The SSDS signal is a reflection of DSB frequency, but it may be influenced by the rate at which DSBs are repaired. To obtain an independent estimate of DSB frequency, we performed

immunostaining for DMC1, a marker for meiotic DSBs, and quantified the distribution of DMC1 foci with respect to telomeres. We compared the density of DMC1 foci in the telomere proximal area with foci density in interstitial regions and found that early in meiosis, when DMC1 foci are first detected (early zygotene), there is a ~1.8-fold excess of DMC1 foci in the telomere-proximal region (fig. S29A). Later in zygotene, although the number of DMC1 foci remains similar [early zygotene:  $185 \pm 37$  (SD); late zygotene:  $166 \pm 41$  (SD);  $P = 0.27$ , two-tailed Mann Whitney], the proportion of telomere-proximal DMC1 foci decreases (Fig. 7, A and B, and figs. S29 and S30). Although we cannot rule out that some subtelomeric DSBs have extended life span, our consistent observation of a decrease in DMC1 foci density near telomeres using different approaches indicates that most of these DSBs do not persist as meiosis progresses. Taken together, our observations suggest that increased DSB formation close to telomeres

in early zygotene shapes the global DSB distribution in our genome-wide maps (Fig. 7C and fig. S31).

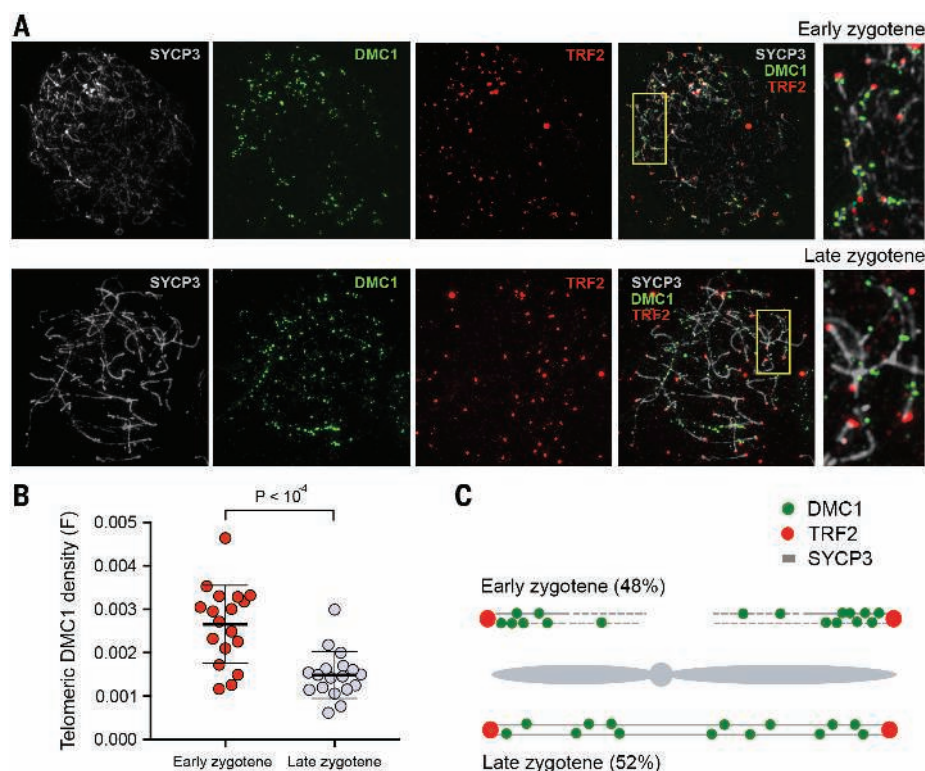
Last, we explored other aspects of DSB formation that may contribute to increased hotspot strength in subtelomeric regions. The PRDM9-dependent H3K4me3 peak intensity from chromatin immunoprecipitation followed by sequencing (ChIP-seq) experiments is correlated with hotspot strength in mice (4, 54); therefore, we examined the H3K4me3 ChIP-seq signal at human hotspots. We found that despite increased hotspot strength in distal regions (figs. S26B, S27, and S32A), the H3K4me3 signal at hotspots did not increase (fig. S32B). Furthermore, the correlation between hotspot strength and H3K4me3 signal decreased in distal relative to interstitial regions (fig. S32C). This suggests that in subtelomeric regions, PRDM9 may define DSBs independent of its methyltransferase activity, or that in these regions, other factors such as the chromatin environment and/or proximity to the nuclear envelope may modulate hotspot strength. Given the biased distribution of crossovers in males but not in females, such modulation may be acting only in the male germ line.

### Meiotic recombination drives genome instability

Meiotic recombination has been implicated as a potential source of gross structural variants (SVs) (55); therefore, we examined the association between meiotic DSBs and SVs. We found that SVs generated by homology-based mechanisms [non-allelic homologous recombination (NAHR) or shrinking or expansion of variable number of tandem repeats (VNTR)] (55) were enriched at *PRDM9*<sub>A</sub>-defined hotspots (table S6). Structural rearrangements derived from unequal crossovers are known to cause genomic diseases (56); therefore, we asked whether disease-causing SV breakpoints occur at DSB hotspots. We found that 14 out of 27 disease-associated breakpoints that have been mapped to <1.5 kb coincided with a *PRDM9*<sub>A</sub>-defined hotspot (table S7). These hotspot-associated breakpoints include those responsible for X-linked ichthyosis (fig. S33A), Charcot-Marie-Tooth disease, and Hunter and Potocki-Lupski/Smith-Magenis syndromes (fig. S33B), among others (table S7). No disease-causing SV breakpoints coincided with *PRDM9*<sub>C</sub>-defined DSB hotspots. This implies that individuals homozygous for the *PRDM9*<sub>C</sub> allele are not at risk for these and highlights the utility of *PRDM9* genotyping for future studies of genomic disorders. Most of these disease-associated breakpoints occur within directly paralogous low copy repeats (DP-LCRs) or inverted paralogous LCRs (IP-LCRs), genomic regions that are susceptible to NAHR-mediated recombination events (57, 58). The 726 *PRDM9*<sub>A</sub>-defined hotspots that occur at DP/IP-LCR regions (file S1) represent targets for future research into human genetic disease.

### Conclusions

We have generated comprehensive maps of meiotic recombination initiation in human individuals. Our comparison of recombination initiation maps between individuals sharing *PRDM9* alleles clearly



**Fig. 7. Increased frequency of DSB formation near telomeres.** (A) Early and late zygotene spermatocytes stained with SYCP3 (which detects axial elements) in gray, DMC1 (a marker of DSBs) in green, and TRF2 (a marker of telomeres) in red. DMC1 foci are clustered near telomeres at early zygotene. (B) Telomere-proximal DMC1 density is about twofold higher in early as compared with late zygotene cells. For each cell, we manually traced all chromosome axes that unambiguously initiated at a TRF2 focus. We defined telomere-proximal regions as the 1  $\mu$ m of axis adjacent to the TRF2 focus (fig. S30A). For each cell, the telomeric density of DMC1 foci (F) was calculated as  $[(\sum \text{DMC1 foci in telomere proximal } 1\text{-}\mu\text{m regions}) / (\text{total DMC1 foci})] / (\text{total TRF2 foci})$ . We do not count DMC1 foci adjacent to TRF2 foci that could not be unambiguously attributed to a particular chromosomal axis. This is particularly punitive for early cells, as the DMC1 density appears highest in regions where TRF2 foci are clustered and individual axes are difficult to distinguish. Error bars indicate mean  $\pm$  SD. P-values are calculated using a Mann-Whitney U-test. (C) Cartoon showing the distribution of DMC1 foci in early and late zygotene spermatocytes. Forty-eight percent of zygotene cells are at the early stage, where they show significantly higher frequency of DSB formation near telomeres. At late zygotene, DSBs are more evenly distributed. The combined signal from these two populations may result in the telomeric bias we observe in our genome-wide maps.



demonstrates that the recombination initiation frequency varies between individuals on a hotspot level. We can explain less than half of the variation in hotspot intensity by sequence changes at PRDM9 binding sites. This suggests that the chromatin environment or other factors mediate the rate of recombination initiation. Our maps allowed us to deduce genome-wide sets of human hotspots defined by different alleles of the PRDM9 protein. A comparison of DSB hotspot maps with LD-based map suggests that an LD-based recombination map is an apparent superposition of allele-specific maps and indicates that a sizable proportion of LD-defined hotspots are defined by minor PRDM9 alleles. Unlike previous methods that rely on the detection of only crossovers, we mapped hotspots by directly identifying the sites of early DSB repair intermediates. Our observations indicate that the DSB frequency itself is largely shaping the crossover distribution.

The high resolution of DSB hotspot mapping allowed us to carefully evaluate the impact of meiotic recombination on genome evolution. We have found clear evidence for GC-biased gene conversion and recombination-associated mutagenesis at sites of DSB hotspots. SVs associated with genomic disease are clearly associated with DSB hotspots defined by the *PRDM9<sub>A</sub>* but not by the *PRDM9<sub>C</sub>* allele, suggesting that the PRDM9 genotype should be considered in assessing predisposition to genomic disorders. Taken together, our data open a broad window into future studies of human recombination, meiosis, and genome evolution and provide a rich data source for future research into human genomic disease.

## Methods Summary

We performed SSDS (27) using antibodies against DMC1 (Santa Cruz Biotechnology, Dallas, TX; C-20, sc 8973) in order to identify meiotic DSB hotspots in testis tissue from five individual human males. The genotype at the *PRDM9* locus was established in each of these individuals by using the primers and method described in (9). SSDS samples were compared with a matched control, and DSB hotspots were defined by using MACS 2.0.10 (59). Whole-genome sequencing libraries were prepared for three individuals according to an established protocol (Illumina, San Diego, CA). GATK best practices (60) were followed to identify variants using the GATK Genome Analysis Toolkit v2.3.9 (61). ChIP using antibodies against H3K4me3 (Abcam; ab8580) followed by high-throughput sequencing was performed to identify sites of trimethylated H3K4 in one individual. Peaks in H3K4me3 data were called by using SICER v1.1 (62). All high-throughput sequencing was performed on an Illumina HiSeq 2500. Spermatoocyte spreads were prepared for immunofluorescence microscopy by using the method described in (63). Detailed methods are available in (33).

## REFERENCES AND NOTES

1. F. Baudat et al., PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327**, 836–840 (2010). doi: [10.1126/science.1183439](https://doi.org/10.1126/science.1183439); pmid: 20044539
2. E. D. Parvanov, P. M. Petkov, K. Paigen, Prdm9 controls activation of mammalian recombination hotspots. *Science* **327**, 835 (2010). doi: [10.1126/science.1181495](https://doi.org/10.1126/science.1181495); pmid: 20044538
3. S. Myers et al., Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science* **327**, 876–879 (2010). doi: [10.1126/science.1182363](https://doi.org/10.1126/science.1182363); pmid: 20044541
4. K. Brick, F. Smagulova, P. Khil, R. D. Camerini-Otero, G. V. Petukhova, Genetic recombination is directed away from functional genomic elements in mice. *Nature* **485**, 642–645 (2012). doi: [10.1038/nature11089](https://doi.org/10.1038/nature11089); pmid: 22660327
5. F. Baudat, Y. Imai, B. de Massy, Meiotic recombination in mammals: Localization and regulation. *Nat. Rev. Genet.* **14**, 794–806 (2013). doi: [10.1038/nrg3573](https://doi.org/10.1038/nrg3573); pmid: 24136506
6. B. de Massy, Initiation of meiotic recombination: How and where? Conservation and specificities among eukaryotes. *Annu. Rev. Genet.* **47**, 563–599 (2013). doi: [10.1146/annurev-genet-110711-155423](https://doi.org/10.1146/annurev-genet-110711-155423); pmid: 24050176
7. P. L. Oliver et al., Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa. *PLoS Genet.* **5**, e1000753 (2009). doi: [10.1371/journal.pgen.1000753](https://doi.org/10.1371/journal.pgen.1000753); pmid: 19997497
8. I. L. Berg et al., Variants of the protein PRDM9 differentially regulate a set of human meiotic recombination hotspots highly active in African populations. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12378–12383 (2011). doi: [10.1073/pnas.1109531108](https://doi.org/10.1073/pnas.1109531108); pmid: 21750151
9. I. L. Berg et al., PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nat. Genet.* **42**, 859–863 (2010). doi: [10.1038/ng.658](https://doi.org/10.1038/ng.658); pmid: 20818382
10. M. J. Neale, S. Keeney, Clarifying the mechanics of DNA strand exchange in meiotic recombination. *Nature* **442**, 153–158 (2006). doi: [10.1038/nature04885](https://doi.org/10.1038/nature04885); pmid: 16838012
11. F. Baudat, B. de Massy, Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis. *Chromosome Res.* **15**, 565–577 (2007). doi: [10.1007/s10577-007-1140-3](https://doi.org/10.1007/s10577-007-1140-3); pmid: 17674146
12. G. A. McVean et al., The fine-scale structure of recombination rate variation in the human genome. *Science* **304**, 581–584 (2004). doi: [10.1126/science.1092500](https://doi.org/10.1126/science.1092500); pmid: 15105499
13. S. Myers, L. Bottolo, C. Freeman, G. McVean, P. Donnelly, A fine-scale map of recombination rates and hotspots across the human genome. *Science* **310**, 321–324 (2005). doi: [10.1126/science.1117196](https://doi.org/10.1126/science.1117196); pmid: 16224025
14. K. A. Frazer et al., A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007). doi: [10.1038/nature06258](https://doi.org/10.1038/nature06258); pmid: 17943122
15. G. R. Abecasis et al., A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010). doi: [10.1038/nature09534](https://doi.org/10.1038/nature09534); pmid: 20981092
16. A. Kong et al., A high-resolution recombination map of the human genome. *Nat. Genet.* **31**, 241–247 (2002). pmid: 12053178
17. A. Kong et al., Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* **467**, 1099–1103 (2010). doi: [10.1038/nature09525](https://doi.org/10.1038/nature09525); pmid: 20981099
18. G. Coop, X. Wen, C. Ober, J. K. Pritchard, M. Przeworski, High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science* **319**, 1395–1398 (2008). doi: [10.1126/science.1151851](https://doi.org/10.1126/science.1151851); pmid: 18239090
19. P. P. Khil, R. D. Camerini-Otero, Genetic crossovers are predicted accurately by the computed human recombination map. *PLOS Genet.* **6**, e1000831 (2010). doi: [10.1371/journal.pgen.1000831](https://doi.org/10.1371/journal.pgen.1000831); pmid: 20126534
20. A. G. Hinch et al., The landscape of recombination in African Americans. *Nature* **476**, 170–175 (2011). doi: [10.1038/nature10336](https://doi.org/10.1038/nature10336); pmid: 21775986
21. N. Arnheim, P. Calabrese, M. Nordborg, Hot and cold spots of recombination in the human genome: The reason we should find them and how this can be achieved. *Am. J. Hum. Genet.* **73**, 5–16 (2003). doi: [10.1086/376419](https://doi.org/10.1086/376419); pmid: 12772086
22. S. Lu et al., Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science* **338**, 1627–1630 (2012). doi: [10.1126/science.1229112](https://doi.org/10.1126/science.1229112); pmid: 23258895
23. J. Wang, H. C. Fan, B. Behr, S. R. Quake, Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell* **150**, 402–412 (2012). doi: [10.1016/j.cell.2012.06.030](https://doi.org/10.1016/j.cell.2012.06.030); pmid: 22817899
24. Y. Hou et al., Genome analyses of single human oocytes. *Cell* **155**, 1492–1506 (2013). doi: [10.1016/j.cell.2013.11.040](https://doi.org/10.1016/j.cell.2013.11.040); pmid: 23460273
25. E. F. Kirkness et al., Sequencing of isolated sperm cells for direct haplotyping of a human genome. *Genome Res.* **23**, 826–832 (2013). doi: [10.1101/gr.144600.112](https://doi.org/10.1101/gr.144600.112); pmid: 23282328
26. F. Cole, S. Keeney, M. Jasin, Comprehensive, fine-scale dissection of homologous recombination outcomes at a hot spot in mouse meiosis. *Mol. Cell* **39**, 700–710 (2010). doi: [10.1016/j.molcel.2010.08.017](https://doi.org/10.1016/j.molcel.2010.08.017); pmid: 20832722
27. P. P. Khil, F. Smagulova, K. M. Brick, R. D. Camerini-Otero, G. V. Petukhova, Sensitive mapping of recombination hotspots using sequencing-based detection of ssDNA. *Genome Res.* **22**, 957–965 (2012). doi: [10.1101/gr.130583.111](https://doi.org/10.1101/gr.130583.111); pmid: 22367190
28. M. I. Jensen-Seaman et al., Comparative recombination rates in the rat, mouse, and human genomes. *Genome Res.* **14**, 528–538 (2004). doi: [10.1101/gr.1970304](https://doi.org/10.1101/gr.1970304); pmid: 15059993
29. D. Thacker, N. Mohibullah, X. Zhu, S. Keeney, Homologue engagement controls meiotic DNA break number and distribution. *Nature* **510**, 241–246 (2014). doi: [10.1038/nature13120](https://doi.org/10.1038/nature13120); pmid: 24717437
30. M. Oliver-Bonet, P. J. Turek, F. Sun, E. Ko, R. H. Martin, Temporal progression of recombination in human males. *Mol. Hum. Reprod.* **11**, 517–522 (2005). doi: [10.1093/molehr/gah193](https://doi.org/10.1093/molehr/gah193); pmid: 16123081
31. P. B. Moens et al., Rad51 immunocytology in rat and mouse spermatocytes and oocytes. *Chromosoma* **106**, 207–215 (1997). doi: [10.1007/s004120050241](https://doi.org/10.1007/s004120050241); pmid: 9254722
32. A. J. Webb, I. L. Berg, A. Jeffreys, Sperm cross-over activity in regions of the human genome showing extreme breakdown of marker association. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10471–10476 (2008). doi: [10.1073/pnas.0804933105](https://doi.org/10.1073/pnas.0804933105); pmid: 18650392
33. Materials and methods are available as supplementary materials on Science Online.
34. K. Hayashi, K. Yoshida, Y. Matsui, A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. *Nature* **438**, 374–378 (2005). doi: [10.1038/nature04112](https://doi.org/10.1038/nature04112); pmid: 16292313
35. S. Myers, C. Freeman, A. Auton, P. Donnelly, G. McVean, A common sequence motif associated with recombination hot spots and genome instability in humans. *Nat. Genet.* **40**, 1124–1129 (2008). doi: [10.1038/ng.213](https://doi.org/10.1038/ng.213); pmid: 19165926
36. T. E. Reddy et al., Effects of sequence variation on differential allelic transcription factor occupancy and gene expression. *Genome Res.* **22**, 860–869 (2012). doi: [10.1101/gr.131201.111](https://doi.org/10.1101/gr.131201.111); pmid: 22300769
37. M. Kasowski et al., Variation in transcription factor binding among humans. *Science* **328**, 232–235 (2010). doi: [10.1126/science.1183621](https://doi.org/10.1126/science.1183621); pmid: 20299548
38. I. Tiemann-Boege, P. Calabrese, D. M. Cochran, R. Sokol, N. Arnheim, High-resolution recombination patterns in a region of human chromosome 21 measured by sperm typing. *PLOS Genet.* **2**, e70 (2006). doi: [10.1371/journal.pgen.0020070](https://doi.org/10.1371/journal.pgen.0020070); pmid: 16680198
39. L. Xu, N. Kleckner, Sequence non-specific double-strand breaks and interhomolog interactions prior to double-strand break formation at a meiotic recombination hot spot in yeast. *EMBO J.* **14**, 5115–5128 (1995). pmid: 7588640
40. T. C. Wu, M. Lichten, Factors that affect the location and frequency of meiosis-induced double-strand breaks in *Saccharomyces cerevisiae*. *Genetics* **140**, 55–66 (1995). pmid: 7635308
41. V. Rocco, A. Nicolas, Sensing of DNA non-homology lowers the initiation of meiotic recombination in yeast. *Genes Cells* **1**, 645–661 (1996). doi: [10.1046/j.1365-2443.1996.00256.x](https://doi.org/10.1046/j.1365-2443.1996.00256.x); pmid: 9078391
42. Q. Q. Fan, F. Xu, M. A. White, T. D. Petes, Competition between adjacent meiotic recombination hotspots in the yeast *Saccharomyces cerevisiae*. *Genetics* **145**, 661–670 (1997). pmid: 9055076
43. J. S. Iacovoni et al., High-resolution profiling of γH2AX around DNA double strand breaks in the mammalian genome. *EMBO J.* **29**, 1446–1457 (2010). doi: [10.1038/emboj.2010.38](https://doi.org/10.1038/emboj.2010.38); pmid: 20360682
44. B. Bost, R. A. Veitia, Dominance and interloci interactions in transcriptional activation cascades: Models explaining compensatory mutations and inheritance patterns. *BioEssays* **36**, 84–92 (2014). doi: [10.1002/bies.201300109](https://doi.org/10.1002/bies.201300109); pmid: 24242332
45. M. W. Nachman, Single nucleotide polymorphisms and recombination rate in humans. *Trends Genet.* **17**, 481–485 (2001). doi: [10.1016/S0168-9525\(01\)00240-X](https://doi.org/10.1016/S0168-9525(01)00240-X); pmid: 11525814

46. C. C. Spencer *et al.*, The influence of recombination on human genetic diversity. *PLOS Genet.* **2**, e148 (2006). doi: [10.1371/journal.pgen.0020148](https://doi.org/10.1371/journal.pgen.0020148); pmid: [17044736](https://pubmed.ncbi.nlm.nih.gov/17044736/)
47. A. Auton *et al.*, A fine-scale chimpanzee genetic map from population sequencing. *Science* **336**, 193–198 (2012). doi: [10.1126/science.1216872](https://doi.org/10.1126/science.1216872); pmid: [22422862](https://pubmed.ncbi.nlm.nih.gov/22422862/)
48. M. Kimura, T. Ohta, The age of a neutral mutant persisting in a finite population. *Genetics* **75**, 199–212 (1973). pmid: [4762875](https://pubmed.ncbi.nlm.nih.gov/4762875/)
49. L. Duret, N. Galtier, Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu. Rev. Genomics Hum. Genet.* **10**, 285–311 (2009). doi: [10.1146/annurev-genom-082908-150001](https://doi.org/10.1146/annurev-genom-082908-150001); pmid: [19630562](https://pubmed.ncbi.nlm.nih.gov/19630562/)
50. Y. Clément, P. F. Arndt, Meiotic recombination strongly influences GC-content evolution in short regions in the mouse genome. *Mol. Biol. Evol.* **30**, 2612–2618 (2013). doi: [10.1093/molbev/mst154](https://doi.org/10.1093/molbev/mst154); pmid: [24030552](https://pubmed.ncbi.nlm.nih.gov/24030552/)
51. M. Ehrlich, R. Y. Wang, 5-Methylcytosine in eukaryotic DNA. *Science* **212**, 1350–1357 (1981). doi: [10.1126/science.6262918](https://doi.org/10.1126/science.6262918); pmid: [6262918](https://pubmed.ncbi.nlm.nih.gov/6262918/)
52. A. Arbel-Eden *et al.*, Trans-lesion DNA polymerases may be involved in yeast meiosis. *G3 (Bethesda)* **3**, 633–644 (2013).
53. A. L. Barlow, M. A. Hultén, Crossing over analysis at pachytene in man. *Eur. J. Hum. Genet.* **6**, 350–358 (1998). doi: [10.1038/sj.ejhg.5200200](https://doi.org/10.1038/sj.ejhg.5200200); pmid: [9781043](https://pubmed.ncbi.nlm.nih.gov/9781043/)
54. F. Smagulova *et al.*, Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. *Nature* **472**, 375–378 (2011). doi: [10.1038/nature09869](https://doi.org/10.1038/nature09869); pmid: [21460839](https://pubmed.ncbi.nlm.nih.gov/21460839/)
55. R. E. Mills *et al.*, Mapping copy number variation by population-scale genome sequencing. *Nature* **470**, 59–65 (2011). doi: [10.1038/nature09708](https://doi.org/10.1038/nature09708); pmid: [21293372](https://pubmed.ncbi.nlm.nih.gov/21293372/)
56. F. Zhang, W. Gu, M. E. Hurles, J. R. Lupski, Copy number variation in human health, disease, and evolution. *Annu. Rev. Genomics Hum. Genet.* **10**, 451–481 (2009). doi: [10.1146/annurev.genom.9.081307.164217](https://doi.org/10.1146/annurev.genom.9.081307.164217); pmid: [19715442](https://pubmed.ncbi.nlm.nih.gov/19715442/)
57. P. Dittwald *et al.*, NAHR-mediated copy-number variants in a clinical population: Mechanistic insights into both genomic disorders and Mendelizing traits. *Genome Res.* **23**, 1395–1409 (2013). doi: [10.1101/gr.152454.112](https://doi.org/10.1101/gr.152454.112); pmid: [23657883](https://pubmed.ncbi.nlm.nih.gov/23657883/)
58. P. Dittwald *et al.*, Inverted low-copy repeats and genome instability—A genome-wide analysis. *Hum. Mutat.* **34**, 210–220 (2013). doi: [10.1002/humu.22217](https://doi.org/10.1002/humu.22217); pmid: [22965494](https://pubmed.ncbi.nlm.nih.gov/22965494/)
59. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008). doi: [10.1186/gb-2008-9-9-r137](https://doi.org/10.1186/gb-2008-9-9-r137); pmid: [18798982](https://pubmed.ncbi.nlm.nih.gov/18798982/)
60. G. Van der Auwera *et al.*, *From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline* (Wiley Online Library, Current Protocols in Bioinformatics, 2013), vol. 43:11.10.1–11.10.33.
61. M. A. DePristo *et al.*, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011). doi: [10.1038/ng.806](https://doi.org/10.1038/ng.806); pmid: [21478889](https://pubmed.ncbi.nlm.nih.gov/21478889/)
62. C. Zang *et al.*, A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952–1958 (2009). doi: [10.1093/bioinformatics/btp340](https://doi.org/10.1093/bioinformatics/btp340); pmid: [19505939](https://pubmed.ncbi.nlm.nih.gov/19505939/)
63. J. Page *et al.*, Inactivation or non-reactivation: What accounts better for the silence of sex chromosomes during mammalian male meiosis? *Chromosoma* **121**, 307–326 (2012). doi: [10.1007/s00412-012-0364-y](https://doi.org/10.1007/s00412-012-0364-y); pmid: [22366883](https://pubmed.ncbi.nlm.nih.gov/22366883/)

## ACKNOWLEDGMENTS

We thank H. Smith and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Genomics Core for sequencing assistance. We thank M. Lichten and P. Hsieh for critical feedback on the manuscript and also extend our gratitude to A. Jeffreys for providing details of human crossover rates. This study used the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD (<http://biowulf.nih.gov>). This research was supported by NIH grant 1R01GM084104-01A1 from National Institute of General Medical Sciences (G.V.P.), March of Dimes grant 1-FY13-506 (G.V.P.) and by the NIDDK Intramural Research Program (R.D.C.O.). The sequencing data reported in this paper are archived at the Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo)) as accession no. GSE59836.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/346/6211/1256442/suppl/DC1](http://www.sciencemag.org/content/346/6211/1256442/suppl/DC1)  
Materials and Methods  
Figs. S1 to S33  
Tables S1 to S7  
References (64–103)  
Data File S1

23 May 2014; accepted 26 September 2014  
[10.1126/science.1256442](https://doi.org/10.1126/science.1256442)



## MESENCHYMAL CELLS

## Defining a mesenchymal progenitor niche at single-cell resolution

Maya E. Kumar,<sup>1</sup> Patrick E. Bogard,<sup>1</sup> F. Hernán Espinoza,<sup>1</sup> Douglas B. Menke,<sup>2,3</sup> David M. Kingsley,<sup>3</sup> Mark A. Krasnow<sup>1\*</sup>

**INTRODUCTION:** In most vertebrate organs, epithelial tubes or sacs are surrounded by support and stromal tissues—including smooth muscle, cartilage, pericytes, fibroblasts, and mesothelium—that form during development from a loose collection of undifferentiated progenitor cells called mesenchyme. Although the behavior and regulation of epithelial progenitors and their niches have begun to be elucidated, much less is known about the identity and behavior of the progenitors of support and stromal tissues. This is critical not only because of the key cell types that they form but also because support and stromal cells can signal to epithelial stem cells and tumors and contribute

to other serious diseases such as fibrosis and asthma. Mesenchyme cells are generally thought to represent highly proliferative, migratory, and multipotent cells that condense around epithelia to generate support and stromal cell types and do not form organized progenitor pools. Elucidating their behavior has been limited by the inability to track the fate of individual mesenchymal cells in development.

**RATIONALE:** We adapted clonal cell labeling strategies with multicolor reporters in mice to probe the behavior and potential of individual and sibling mesenchyme cells in lung development. This was used to define the proliferation, migration, and

differentiation behavior of individual mesenchyme cells and to map the locations and behavior of mesenchymal progenitors at single-cell resolution.

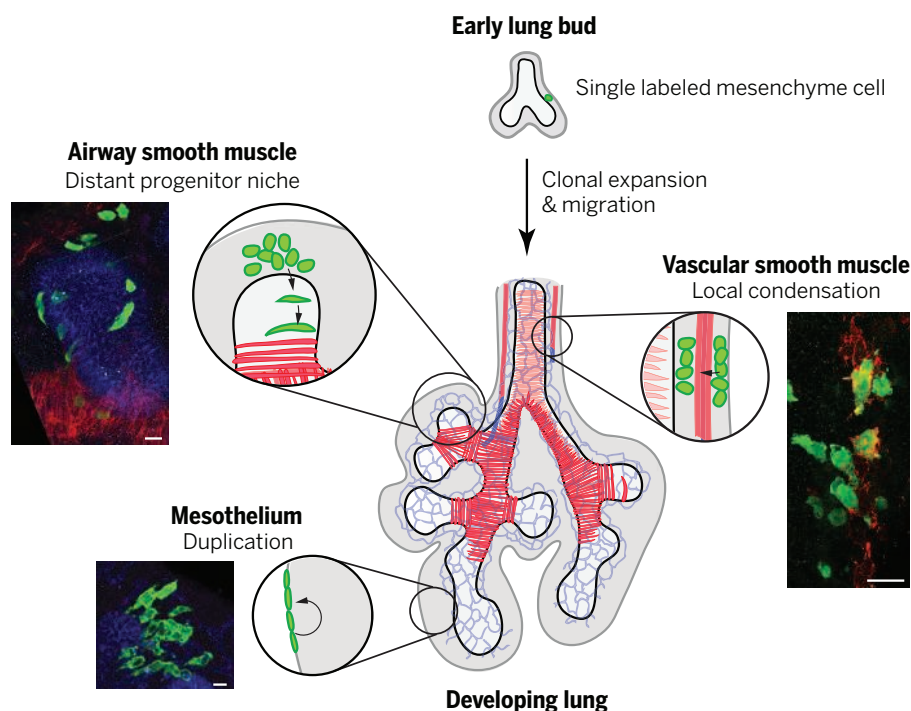
**RESULTS:** We show that although mesenchymal cells are highly proliferative, as classical studies suggested, there is a surprising diversity of mesenchymal progenitor populations with different locations,

## ON OUR WEB SITE

Read the full article at <http://dx.doi.org/10.1126/science.1258810>

patterns of migration, recruitment mechanisms, and lineage boundaries. We focus on airway smooth muscle progenitors, which map exclusively to the mesenchyme just ahead of budding and bifurcating airway branches. Progenitors are recruited from these tip pools to the branch stalk, where they differentiate into circumferentially oriented airway smooth muscle cells. There is a lineage boundary that prevents mesenchymal cells surrounding airway stalks from becoming airway smooth muscle from branch sides, but this stalk mesenchyme can be induced in the presence of a newly budded airway branch to generate a new smooth muscle progenitor pool dedicated to the new branch. Micrografting experiments show that the airway tip alone is not sufficient to induce a new smooth muscle progenitor pool in stalk mesenchyme. The missing mesenchymal signal can be provided by a focal Wnt signal, and delocalized Wnt pathway activity expands or alters the progenitor pool, causing ectopic smooth muscle formation on nearby endothelial cells.

**CONCLUSIONS:** Lung mesenchyme is neither a large homogeneous progenitor pool, as it has classically been viewed, nor a collection of discrete, isolated, and unchanging progenitor niches. Rather, the progenitors of each differentiated cell type occupy different locations and display varied modes of recruitment. The localized and carefully controlled domains of airway smooth muscle progenitors rival epithelial progenitor niches in regulatory sophistication. ■



**Diverse mechanisms generate mesenchymal cell derivatives.** (Top) Single mesenchyme cells are labeled early in lung development. The labeled cell (green) proliferates, and daughter cells disperse to seed progenitor “niches” that generate support and stromal cell types. (Bottom) Photomicrographs of individual clones and schematics of mesenchymal niches, highlighting the distinct modes of recruitment. Airway epithelium, white (schematic), blue (photomicrographs); mesenchyme, light gray (schematic); mesothelium, dark gray (schematic outline); smooth muscle, red (schematic); endothelium, blue (schematic). Scale bars, 10  $\mu$ m.

<sup>1</sup>Howard Hughes Medical Institute and Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305-5307, USA. <sup>2</sup>Department of Genetics, University of Georgia, Athens, GA 30602-2607, USA. <sup>3</sup>Howard Hughes Medical Institute and Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305-5329, USA.

\*Corresponding author. E-mail: [krasnow@stanford.edu](mailto:krasnow@stanford.edu)  
Cite this article as M. E. Kumar *et al.*, *Science* **346**, 1258810 (2014). DOI: 10.1126/science.1258810

## RESEARCH ARTICLE

## MESENCHYMAL CELLS

# Defining a mesenchymal progenitor niche at single-cell resolution

Maya E. Kumar,<sup>1</sup> Patrick E. Bogard,<sup>1</sup> F. Hernán Espinoza,<sup>1</sup> Douglas B. Menke,<sup>2,3</sup> David M. Kingsley,<sup>3</sup> Mark A. Krasnow<sup>1\*</sup>

Most vertebrate organs are composed of epithelium surrounded by support and stromal tissues formed from mesenchyme cells, which are not generally thought to form organized progenitor pools. Here, we use clonal cell labeling with multicolor reporters to characterize individual mesenchymal progenitors in the developing mouse lung. We observe a diversity of mesenchymal progenitor populations with different locations, movements, and lineage boundaries. Airway smooth muscle (ASM) progenitors map exclusively to mesenchyme ahead of budding airways. Progenitors recruited from these tip pools differentiate into ASM around airway stalks; flanking stalk mesenchyme can be induced to form an ASM niche by a lateral bud or by an airway tip plus focal Wnt signal. Thus, mesenchymal progenitors can be organized into localized and carefully controlled domains that rival epithelial progenitor niches in regulatory sophistication.

Most vertebrate organs are composed of epithelial tubes or sacs surrounded by support and stromal tissues that form during development from a loose collection of undifferentiated progenitor cells called mesenchyme (1). Mesenchyme gives rise to a variety of cell types—including cartilage, smooth muscle, pericytes, mesothelium, and fibroblasts—whose development and ultimate structural organization are intimately associated with the epithelia they surround (2, 3). Mesenchymal cells have long been of interest to developmental biologists for their role as a source of signals that induce and pattern epithelial cell types (4, 5). More recently, there has been a surge of interest in these cells because of their ability to serve as signaling centers that support stem cells and influence tumor progression (6, 7) and because mesenchyme-derived cell types contribute to many serious diseases, such as fibrosis of the lung and other organs (8), and airway wall remodeling in asthma (9).

Although the rules governing the behavior and regulation of a variety of epithelial progenitor cells and their niches have begun to be elucidated (10, 11), much less is known about the identity and behavior of stromal cells and other progenitors in the mesenchyme (12). The classical view is that mesenchyme cells are highly proliferative, migratory, and multipotent cells that expand during organ development and condense around endodermal epithelium to generate diverse support and stromal cell types (1, 13). Epithelial signals,

including Wnt, fibroblast growth factor (FGF), and Hedgehog family members, have been implicated in these processes (14–16), but how the pathways affect mesenchymal progenitor cells is not well understood due to the variety of signals and challenges in following the behavior of individual mesenchymal cells in vivo.

Here, we adapt clonal cell labeling strategies in mice, like those recently used to elucidate epithelial progenitor and stem cell behavior in vivo (17–19), to probe the behavior and potential of individual mesenchyme cells during lung development. The developing lung is composed of a branching airway tree and corresponding networks of developing blood vessels, all surrounded by undifferentiated pulmonary mesenchyme cells. Each of these tubular networks is associated with distinct smooth muscle support layers and other intervening cell types formed from the mesenchyme (20). Our clonal analysis shows that mesenchymal progenitors are highly proliferative, as classical studies suggest, and they are also highly motile. However, their motility is not random, and we show how clonal analysis can be used to map mesenchymal progenitor niches and progenitor cell behavior at cellular resolution. This reveals a surprising diversity of mesenchymal progenitor populations with different locations, patterns of movement, and migration and lineage boundaries. We focus on the airway smooth muscle (ASM) progenitor population, which displays a remarkable degree of regulation that rivals that of epithelial stem cells, and show that a focal Wnt signal can regulate movement of ASM progenitors out of their niche.

## Results

### Marking cell clones in the developing lung mesenchyme

To identify progenitor cells in mouse lung mesenchyme and characterize their behavior at single-

cell resolution, we carried out a clonal analysis in which individual lung mesenchyme cells were genetically marked and the structures of the resulting clones analyzed to reveal the proliferation, migration, and developmental potential of individual progenitor cells (Fig. 1, A and B). Two Cre recombinase-dependent clonal labeling strategies were used to mark individual lung mesenchyme cells. In one, a ubiquitously expressed hypoxanthine-guanine phosphoribosyltransferase-Cre (HPRT-Cre) allele was used to catalyze rare interchromosomal recombination events between mosaic analysis with double markers (MADM) chromosomes, which recreate the complete coding sequence for green fluorescent protein (GFP) on one resultant chromosome and the complete coding sequence for *Drosophila* sp. red fluorescent protein (DsRed) on the other (21). This generates either red and green twin-spot clones or yellow clones in all major lineages of the developing mouse lung (Fig. 1, C, E, and F), and we focused on clones in the mesenchyme and its derivatives. Although HPRT-Cre is expressed from the earliest stages of embryogenesis (22), cell labeling in the lung was first observed at embryonic day 11.5 (E11.5) and continues thereafter; we restricted our analysis to between E11.5 and E13.5, when individual clones could be resolved. All MADM samples were immunostained and analyzed in whole mount to preserve the three-dimensional arrangement of cells in the clone and their relationship to the rest of the developing lung.

To gain temporal and spatial control of clone generation, we combined a lung-specific mesenchyme enhancer element from the mouse *Tbx4* locus (23) with an Hsp68 minimal promoter to drive Cre expression (*Tbx4*<sup>LME</sup>-Cre) (Fig. 1G) exclusively in the lung mesenchyme of transgenic mice (Fig. 1K). Cre mRNA expression is robustly detected in early lung buds by E10.5 and continues through E14.5 but is undetectable by birth (Fig. 1, H to J). When crossed to a Cre reporter, recombination is observed in all lineages of the lung mesenchyme but excluded from both epithelium (Fig. 1L) and endothelium (24). Virtually all lung mesenchyme and its derivatives are labeled by *Tbx4*<sup>LME</sup>-Cre by E13.5, including undifferentiated mesenchyme, ASM, vascular smooth muscle (VSM), and mesothelium. This Cre mouse line thus allows targeted labeling and other genetic manipulations specifically in lung mesenchyme.

We also constructed a *Tbx4*<sup>LME</sup>-CreER transgenic mouse (fig. S1) in which Cre is fused to a modified estrogen receptor to allow temporal control of clone marking and the extent of cell labeling. We used the CreER transgene with the multicolor Rainbow reporter, in which an individual recombination event is marked by expression of one of three possible fluorescent proteins (25). We titrated the dose of tamoxifen such that just one, or at most two, recombination events occurred in the lung, indicated by the presence of marked cells of a single color or two different colors (Fig. 1, M to O). Lungs in which more than two colors were detected were excluded from analysis, and all cells of the same color within a lung were considered a clone. Because of the uncertainty about when MADM

<sup>1</sup>Howard Hughes Medical Institute and Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305-5307, USA. <sup>2</sup>Department of Genetics, University of Georgia, Athens, GA 30602-2607, USA. <sup>3</sup>Howard Hughes Medical Institute and Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305-5329, USA.

\*Corresponding author. E-mail: krasnow@stanford.edu



clones were induced, they are not included in the major clone table (Table 1); however, below we use examples of clones generated with both the Rainbow and MADM clone-marking systems (and separately tabulate the MADM examples) because our findings were similar for both.

### Clonal analysis reveals cell dynamics of developing mesenchyme

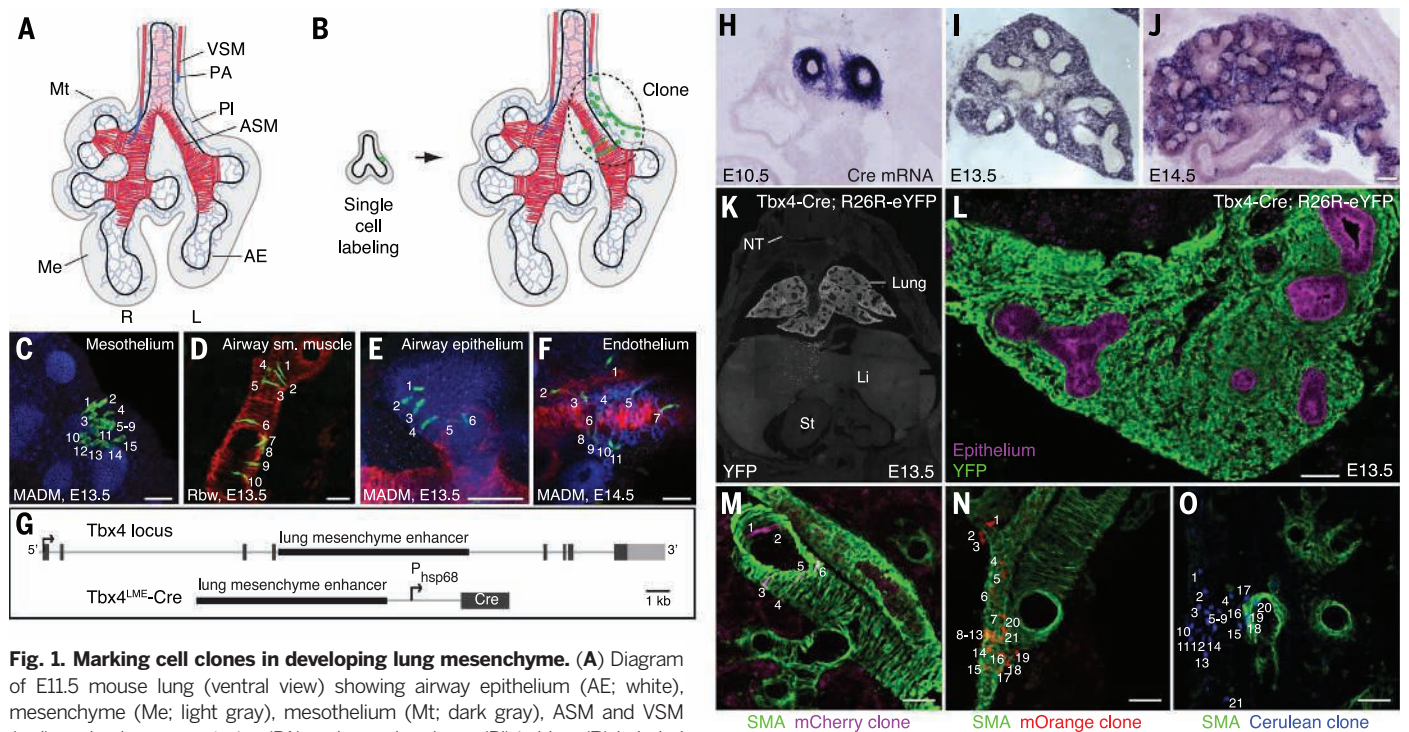
Principles governing cell behavior in the developing mesenchyme can be deduced by examining random clones (26). Comparison of cell number in each clone showed that proliferation rates vary substantially among mesenchyme cells. Rainbow clones induced at E10.5 and analyzed at E13.5 showed a range of 4 to 134 cells, giving calculated doubling times of 8 to 45 hours, assuming a 12-hour lag between tamoxifen administration and recombination (Table 1). Comparison of the E13.5 clones with ones analyzed at E14.5 identified a striking example of proliferative regulation (Fig. 2A). At E13.5, all clones, regardless of chase time or degree of differentiation, were composed of fewer than 200 cells, whereas at E14.5 a second class of clones was also observed (boxed

in Fig. 2A) that contained many more cells (257 to 840 cells) with various differentiated mesenchymal cells but dominated by undifferentiated mesenchyme (fibroblasts) (Table 1 and fig. S2A). Simulations indicate that these large, fibroblast-dominated clones are unlikely to be misassigned polyclones (fig. S2B) but rather identify a subpopulation of highly proliferative fibroblasts that arise just after E13.5.

The MADM labeling system allows comparison of the behavior of sister cells in twinspace clones, red or green marked clones generated from a single parent cell (Fig. 2, C to F). Such twinspace clones showed frequent and, in some cases, dramatic asymmetries in the number of cells generated from each sister cell (Fig. 2, D and E). The ratio of cell number in twinspace clones ranged from 1 to 11 ( $n = 53$  twinspace clones; table S1), with 68% of twinspace clones differing by a factor of more than 1.5 (Fig. 2C) and, in extreme cases ( $n = 7$  of 53), the complete absence of a red or green twinspace (Fig. 2E). Simulations indicated that such differences in sibling cell proliferation rates are unlikely to result solely from stochastic variation (Fig. 2C) ( $P = 3 \times 10^{-101}$  versus Poisson simulation), im-

plying important biological regulation of sister cell proliferation and/or survival.

Clonal analysis also revealed a large amount of migration and intermixing among cells of the mesenchyme. Only 26% of labeled cells ( $n = 575$  cells scored from five randomly selected clones) remained in contact with a sibling. These dispersed clones spanned as much as 1700  $\mu\text{m}$  along the proximal-distal axis over a chase period of four days (Table 1). Although 81% of clones ( $n = 70$ ) were confined to a single lung lobe, clones induced as late as E11.5 were able to contribute to multiple lobes (Table 1), implying that divisions between lobes are not complete until after that stage. However, clones spanning left and right lungs were never observed, even among clones induced at E9.5 when lung buds first emerge, indicating that the division between left and right is established early and cell movement between them occurs rarely or not at all (Table 1). MADM samples analyzed in whole mount showed that most mesenchyme clones were found in cell clouds of no regular shape (Fig. 2D), but when clones showed a directional bias, cells were arrayed lengthwise along airway



**Fig. 1. Marking cell clones in developing lung mesenchyme.** (A) Diagram of E11.5 mouse lung (ventral view) showing airway epithelium (AE; white), mesenchyme (Me; light gray), mesothelium (Mt; dark gray), ASM and VSM (red), and pulmonary arteries (PA) and vascular plexus (PI) in blue. (B) Labeled mesenchymal progenitor cell (green) in early lung (left), which at E11.5 (right) has generated a mixed lineage clone (green, dashed circle) with labeled ASM, mesothelial cells, and undifferentiated fibroblasts. (C to F). Close-up of clones (pseudocolored green) in indicated tissues of lungs of indicated ages. Clones were marked with MADM [(C), (E), and (F)] or Rainbow (D) strategies and immunostained for E-cadherin (airway epithelium, blue); platelet endothelial cell adhesion molecule (PECAM) [vascular endothelium, blue in (F)]; and smooth muscle  $\alpha$ -actin (ASM and VSM, red). Cells in each clone are numbered. (G) Structure of lung mesenchyme-specific Cre transgene using lung mesenchyme-specific enhancer (LME) from *Tbx4* locus and  $P_{hsp68}$  minimal promoter to drive Cre recombinase expression. (H to J) Cre mRNA in situ hybridization of lungs from *Tbx4*<sup>LME</sup>-Cre transgenic mice of indicated ages. mRNA is robustly detected by E10.5 exclusively in lung mesenchyme (H), remains strong at E13.5 (I), and

begins to down-regulate by E14.5 (J). Similar results were obtained for *Tbx4*<sup>LME</sup>-CreER (fig. S1). (K) Tiled image of transverse section through E13.5 *Tbx4*<sup>LME</sup>-Cre;R26R-EYFP lung immunostained for yellow fluorescent protein (YFP) lineage trace marker (white). Lineage-traced cells are confined to lung. NT, neural tube; Li, liver; St, stomach. (L) Section through E13.5 *Tbx4*<sup>LME</sup>-Cre;R26R-EYFP lung immunostained for YFP (green) and E-cadherin (magenta). Nearly all mesenchyme cells are lineage labeled (green). Lineage label is not detected in epithelium (magenta). (M to O) *Tbx4*<sup>LME</sup>-CreER;Rainbow embryos given limiting doses of tamoxifen (0.6 mg per pregnant female) at E9.5 (M) or E10.5 [(N) and (O)] and analyzed at E13.5 by immunostaining for smooth muscle  $\alpha$ -actin (SMA) and clone markers indicated. Clones are labeled by a single recombination reporter, either mCherry (M), mOrange (N), or Cerulean (O). Cells in each clone are numbered. Scale bars, 50  $\mu\text{m}$  [(C) to (F) and (H) to (J)], 100  $\mu\text{m}$  [(L) to (O)].

Table 1. Tbx4<sup>LME</sup>-CreERT2;Rainbow mesenchyme clones<sup>1</sup>

| Clone no. | Lobe    | Chase        |                      | No. cells | Span (μm) | Est. hrs/div <sup>3</sup> | % of individual cell types |     |     |     |     |                |   |   |   |   |     |   | % diff <sup>4</sup> | P-value <sup>5</sup> |
|-----------|---------|--------------|----------------------|-----------|-----------|---------------------------|----------------------------|-----|-----|-----|-----|----------------|---|---|---|---|-----|---|---------------------|----------------------|
|           |         | Tamox (days) | Harvest <sup>2</sup> |           |           |                           | F                          | A   | V   | M   | C   | F <sup>4</sup> | A | V | M | C |     |   |                     |                      |
| 1         | n.d.    | E11.5        | 3                    | E14.5     | 3         | 50                        | 38-42                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 0.90                |                      |
| 2         | n.d.    | E9.5         | 4                    | E13.5     | 4         | 50                        | 42-45                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 0.85                |                      |
| 3         | n.d.    | E11.5        | 3                    | E14.5     | 9         | 50                        | 19-21                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 0.54                |                      |
| 4         | RAc     | E9.5         | 4                    | E13.5     | 11        | 150                       | 24-26                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 0.43                |                      |
| 5         | RCd     | E10.5        | 3                    | E13.5     | 14        | 100                       | 16-17                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 0.30                |                      |
| 6         | R       | E10.5        | 4                    | E14.5     | 15        | 100                       | 22-23                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 0.27                |                      |
| 7         | RCd     | E11.5        | 2                    | E13.5     | 21        | 50                        | 8-10                       | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 0.12                |                      |
| 8         | Rmd     | E10.5        | 4                    | E14.5     | 23        | n.d.                      | 19-20                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 0.09                |                      |
| 9         | RCd     | E10.5        | 3                    | e13.5     | 40        | 150                       | 11-12                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 7.9 E-3 *           |                      |
| 10        | n.d.    | E11.5        | 3                    | E14.5     | 53        | 450                       | 10-12                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 1.1 E-3 *           |                      |
| 11        | L       | E10.5        | 4                    | E14.5     | 61        | 500                       | 14-15                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 3.1 E-4             |                      |
| 12        | n.d.    | E11.5        | 3                    | E14.5     | 61        | n.d.                      | 10-11                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 3.1 E-4             |                      |
| 13        | RCd     | E10.5        | 4                    | E14.5     | 96        | 200                       | 13-14                      | 100 | 0   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 0   | 0 | 1.1 E-6 *           |                      |
| 14        | RCd     | E11.5        | 2                    | E13.5     | 1         | 50                        | n/a                        | 0   | 100 | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 0.37                |                      |
| 15        | RAc     | E10.5        | 3                    | E13.5     | 6         | 100                       | 23-26                      | 0   | 100 | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 3.4 E-5 *           |                      |
| 16        | Rmd     | E9.5         | 4                    | E13.5     | 6         | 50                        | 33-35                      | 0   | 100 | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 3.4 E-5 *           |                      |
| 17        | RCr     | E10.5        | 3                    | E13.5     | 7         | 60                        | 21-24                      | 0   | 100 | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 4.6 E-6 *           |                      |
| 18        | L       | E11.5        | 2                    | E13.5     | 11        | 500                       | 10-12                      | 0   | 100 | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 1.3 E-9 *           |                      |
| 19        | Tr+L    | E10.5        | 3                    | E13.5     | 12        | 100                       | 17-18                      | 0   | 100 | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 1.6 E-10 *          |                      |
| 20        | L       | E10.5        | 4                    | E14.5     | 12        | 150                       | 23-25                      | 0   | 100 | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 1.6 E-10 *          |                      |
| 21        | Rmd     | E10.5        | 4                    | E14.5     | 13        | 100                       | 23-24                      | 0   | 100 | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 2.0 E-11 *          |                      |
| 22        | Rmd     | E10.5        | 3                    | E13.5     | 11        | 50                        | 17-19                      | 0   | 0   | 100 | 0   | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 8.2 E-70 *          |                      |
| 23        | L       | E10.5        | 4                    | E14.5     | 5         | 50                        | 36-39                      | 0   | 0   | 0   | 100 | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 5.4 E-37 *          |                      |
| 24        | Rmd     | E10.5        | 3                    | E13.5     | 11        | 180                       | 17-19                      | 0   | 0   | 0   | 100 | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 1.7 E-82 *          |                      |
| 25        | Tr      | E10.5        | 3                    | E13.5     | 12        | 50                        | 17-18                      | 0   | 0   | 0   | 100 | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 4.1 E-90 *          |                      |
| 26        | Tr      | E10.5        | 4                    | E14.5     | 17        | 150                       | 21-22                      | 0   | 0   | 0   | 100 | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 3.6 E-128 *         |                      |
| 27        | L       | E9.5         | 4                    | E13.5     | 64        | 450                       | 14-15                      | 0   | 0   | 0   | 100 | 0              | 0 | 0 | 0 | 0 | 100 | 0 | 0                   |                      |
| 28        | L       | E10.5        | 3                    | E13.5     | 4         | 50                        | 30-33                      | 50  | 50  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 50  | 0 | 0.62                |                      |
| 29        | RAc     | E10.5        | 3                    | E13.5     | 5         | 20                        | 26-28                      | 60  | 40  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 40  | 0 | 0.80                |                      |
| 30        | Rmd     | E11.5        | 2                    | E13.5     | 10        | 100                       | 11-13                      | 20  | 80  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 80  | 0 | 6.3 E-5 *           |                      |
| 31        | L       | E10.5        | 3                    | E13.5     | 11        | 100                       | 17-19                      | 36  | 64  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 64  | 0 | 5.6 E-3 *           |                      |
| 32        | L       | E10.5        | 3                    | E13.5     | 14        | 50                        | 16-17                      | 64  | 36  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 36  | 0 | 0.52                |                      |
| 33        | L       | E10.5        | 4                    | E14.5     | 20        | 900                       | 19-21                      | 30  | 70  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 70  | 0 | 6.1 E-7 *           |                      |
| 34        | Rmd     | E10.5        | 4                    | E14.5     | 23        | 100                       | 19-20                      | 83  | 17  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 17  | 0 | 0.77                |                      |
| 35        | Tr      | E10.5        | 4                    | E14.5     | 28        | 600                       | 18-19                      | 50  | 54  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 54  | 0 | 2.0 E-4 *           |                      |
| 36        | Rmd     | E11.5        | 3                    | E14.5     | 29        | 200                       | 12-14                      | 83  | 17  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 17  | 0 | 0.68                |                      |
| 37        | RAc     | E10.5        | 3                    | E13.5     | 30        | 200                       | 12-13                      | 27  | 73  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 73  | 0 | 6.1 E-12 *          |                      |
| 38        | RCr     | E11.5        | 3                    | E14.5     | 32        | 200                       | 12-13                      | 94  | 6   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 6   | 0 | 0.16                |                      |
| 39        | RCr     | E11.5        | 3                    | E14.5     | 41        | 250                       | 11-12                      | 90  | 10  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 10  | 0 | 0.20                |                      |
| 40        | n.d.    | E10.5        | 3                    | E13.5     | 55        | 300                       | 10-11                      | 91  | 9   | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 9   | 0 | 7.3 E-2             |                      |
| 41        | RCd     | E10.5        | 3                    | E13.5     | 57        | 300                       | 10-11                      | 61  | 39  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 39  | 0 | 2.0 E-3 *           |                      |
| 42        | n.d.    | E10.5        | 4                    | E14.5     | 76        | 750                       | 13-14                      | 70  | 30  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 30  | 0 | 3.1 E-2 *           |                      |
| 43        | Acc     | E10.5        | 4                    | E14.5     | 83        | 350                       | 13-14                      | 86  | 14  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 14  | 0 | 9.4 E-2             |                      |
| 44        | Rmd     | E10.5        | 3                    | E13.5     | 92        | 750                       | 9-10                       | 59  | 41  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 41  | 0 | 7.7 E-7 *           |                      |
| 45        | L       | E10.5        | 4                    | E14.5     | 111       | 300                       | 12-13                      | 71  | 29  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 29  | 0 | 8.7 E-3 *           |                      |
| 46        | n.d.    | E10.5        | 3                    | E13.5     | 134       | 250                       | 8-9                        | 67  | 10  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 10  | 0 | 4.5 E-3 *           |                      |
| 47        | n.d.    | E10.5        | 4                    | E14.5     | 139       | 600                       | 12-13                      | 71  | 29  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 29  | 0 | 1.2 E-3 *           |                      |
| 48        | L       | E10.5        | 4                    | E14.5     | 149       | 650                       | 12                         | 74  | 26  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 26  | 0 | 6.0 E-3 *           |                      |
| 49        | RCr+RCd | E10.5        | 4                    | E14.5     | 399       | 1250                      | 10                         | 77  | 23  | 0   | 0   | 0              | 0 | 0 | 0 | 0 | 23  | 0 | 2.5 E-6 *           |                      |
| 50        | R       | E10.5        | 3                    | E13.5     | 10        | 20                        | 18-20                      | 20  | 0   | 80  | 0   | 0              | 0 | 0 | 0 | 0 | 80  | 0 | 1.4 E-39 *          |                      |
| 51        | n.d.    | E11.5        | 3                    | E14.5     | 17        | n.d.                      | 15-16                      | 47  | 0   | 53  | 0   | 0              | 0 | 0 | 0 | 0 | 53  | 0 | 3.2 E-28 *          |                      |
| 52        | R       | E10.5        | 3                    | E13.5     | 21        | 50                        | 14-15                      | 24  | 0   | 76  | 0   | 0              | 0 | 0 | 0 | 0 | 76  | 0 | 6.2 E-76 *          |                      |
| 53        | RCr     | E10.5        | 4                    | E14.5     | 16        | 50                        | 21-23                      | 56  | 0   | 44  | 0   | 0              | 0 | 0 | 0 | 0 | 44  | 0 | 4.4 E-21 *          |                      |
| 54        | n.d.    | E10.5        | 4                    | E14.5     | 28        | 250                       | 17-19                      | 71  | 25  | 4   | 0   | 0              | 0 | 0 | 0 | 0 | 25  | 0 | 0.81                |                      |
| 55        | L       | E10.5        | 3                    | E13.5     | 43        | 90                        | 11-12                      | 33  | 12  | 56  | 0   | 0              | 0 | 0 | 0 | 0 | 67  | 0 | 7.5 E-81 *          |                      |
| 56        | n.d.    | E10.5        | 4                    | E14.5     | 46        | n.d.                      | 15-16                      | 78  | 17  | 4   | 0   | 0              | 0 | 0 | 0 | 0 | 22  | 0 | 0.74                |                      |
| 57        | L       | E11.5        | 3                    | E14.5     | 49        | 150                       | 11-12                      | 80  | 10  | 10  | 0   | 0              | 0 | 0 | 0 | 0 | 20  | 0 | 2.4 E-2 *           |                      |
| 58        | L       | E10.5        | 3                    | E13.5     | 55        | 120                       | 10-11                      | 25  | 60  | 15  | 0   | 0              | 0 | 0 | 0 | 0 | 75  | 0 | 9.1 E-19 *          |                      |
| 59        | n.d.    | E10.5        | 3                    | E13.5     | 56        | 200                       | 10-11                      | 82  | 11  | 7   | 0   | 0              | 0 | 0 | 0 | 0 | 18  | 0 | 0.13                |                      |
| 60        | RMD+RAc | E10.5        | 3                    | E13.5     | 70        | 240                       | 10-11                      | 80  | 17  | 3   | 0   | 0              | 0 | 0 | 0 | 0 | 20  | 0 | 0.56                |                      |
| 61        | All R   | E10.5        | 4                    | E14.5     | 73        | 550                       | 14-15                      | 86  | 8   | 5   | 0   | 0              | 0 | 0 | 0 | 0 | 14  | 0 | 4.9 E-2 *           |                      |
| 62        | Tr+R    | E10.5        | 4                    | E14.5     | 73        | 1700                      | 14-15                      | 16  | 63  | 21  | 0   | 0              | 0 | 0 | 0 | 0 | 84  | 0 | 1.5 E-37 *          |                      |
| 63        | RCr     | E10.5        | 4                    | E14.5     | 93        | 150                       | 13-14                      | 75  | 16  | 9   | 0   | 0              | 0 | 0 | 0 | 0 | 25  | 0 | 1.8 E-2 *           |                      |
| 64        | RCr+R   | E10.5        | 3                    | E13.5     | 94        | 550                       | 9-10                       | 39  | 49  | 12  | 0   | 0              | 0 | 0 | 0 | 0 | 61  | 0 | 1.7 E-17 *          |                      |
| 65        | n.d.    | E10.5        | 3                    | E13.5     | 99        | n.d.                      | 9-10                       | 72  | 25  | 3   | 0   | 0              | 0 | 0 | 0 | 0 | 28  | 0 | 0.21                |                      |
| 66        | All R   | E10.5        | 4                    | E14.5     | 687       | 1350                      | 9-10                       | 90  | 7   | 3   | 0   | 0              | 0 | 0 | 0 | 0 | 10  | 0 | 4.5 E-21 *          |                      |
| 67        | RMD+R   | E10.5        | 4                    | E14.5     | 53        | 800                       | 15-16                      | 38  | 40  | 0   | 0   | 23             | 0 | 0 | 0 | 0 | 62  | 0 | 1.4 E-69 *          |                      |
| 68        | All R   | E10.5        | 4                    | E14.5     | 840       | n.d.                      | 9                          | 83  | 12  | 0   | 0   | 5              | 0 | 0 | 0 | 0 | 17  | 0 | 3.2 E-49 *          |                      |
| 69        | L       | E10.5        | 4                    | E14.5     | 18        | 200                       | 20-22                      | 44  | 6   | 0   | 50  | 0              | 0 | 0 | 0 | 0 | 56  | 0 | 2.1 E-31 *          |                      |
| 70        | n.d.    | E11.5        | 3                    | E14.5     | 31        | n.d.                      | 12-13                      | 81  | 6   | 0   | 13  | 0              | 0 | 0 | 0 | 0 | 19  | 0 | 3.8 E-3 *           |                      |
| 71        | RCr+RCd | E10.5        | 4                    | E14.5     | 105       | 550                       | 13                         | 67  | 21  | 0   | 2   | 0              | 0 | 0 | 0 | 0 | 23  | 0 | 0.23                |                      |
| 72        | RMD+RCd | E10.5        | 4                    | E14.5     | 137       | 1600                      | 12-13                      | 75  | 19  | 0   | 6   | 0              | 0 | 0 | 0 | 0 | 25  | 0 | 3.6 E-2 *           |                      |
| 73        | n.d.    | E10.5        | 4                    | E14.5     | 710       | n.d.                      | 9-10                       | 82  | 18  | 0   | 1   | 0              | 0 | 0 | 0 | 0 | 18  | 0 | 6.5 E-9 *           |                      |
| 74        | RAc     | E10.5        | 4                    | E14.5     | 101       | 600                       | 13-14                      | 44  | 27  | 27  | 3   | 0              | 0 | 0 | 0 | 0 | 56  | 0 | 8.7 E-40 *          |                      |
| 75        | n.d.    | E11.5        | 3                    | E14.5     | 112       | n.d.                      | 9-10                       | 77  | 14  | 4   | 5   | 0              | 0 | 0 | 0 | 0 | 23  | 0 | 0.28                |                      |
| 76        | Rmd     | E10.5        | 4                    | E14.5     | 143       | 650                       | 12-13                      | 64  | 29  | 3   | 4   | 0              | 0 | 0 | 0 | 0 | 36  | 0 | 2.1 E-2 *           |                      |
| 77        | L       | E10.5        | 4                    | E14.5     | 257       | 800                       | 10-11                      | 74  | 14  | 9   | 3   | 0              | 0 | 0 | 0 | 0 | 26  | 0 | 2.1 E-6 *           |                      |
| 78        | All R   | E11.5        | 3                    | E14.5     | 301       | 1350                      | 7-8                        | 87  | 8   | 1   | 4   | 1              | 0 | 0 | 0 | 0 | 13  | 0 | 1.5 E-6 *           |                      |

<sup>1</sup>Abbreviations: L, left lobe; n.d., not determined; RAc, accessory lobe; RCd, right caudal lobe; RCr, right cranial lobe; RMD, right middle lobe; all R, all right lobes; F, fibroblast (undifferentiated); A, airway smooth muscle; V, vascular smooth muscle; M, mesothelium; C, cartilage

<sup>2</sup>Clones highlighted in gray were collected at E14.5

<sup>3</sup>Estimated doubling times assuming recombination 6-12 hours after tamoxifen administration, uniform division among clone cells, and no cell death

<sup>4</sup>Bar graph showing cell type composition of each clone, with most abundant cell type highlighted in red

<sup>5</sup>Percent of total cells in clone that scored as one of the four differentiated cell types

<sup>6</sup>Likelihood the cell type composition of a clone arose



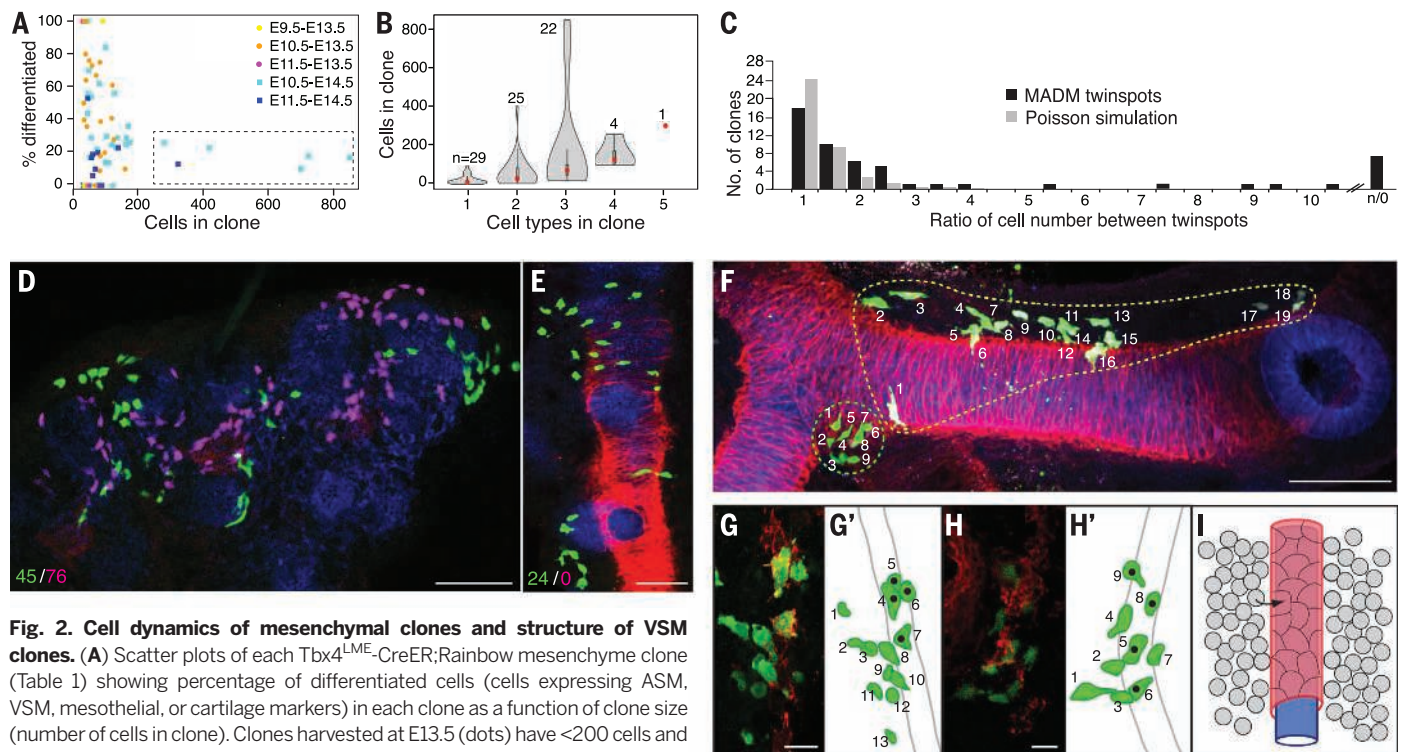
cells to form VSM from mesenchyme immediately surrounding the vessel (Fig. 2I) (30). Below, we use this approach to map progenitors of ASM.

### ASM progenitors reside at branch tips

Double labeling with cell type-specific antibodies to highlight both the airway epithelium and developing smooth muscle (Fig. 3A) shows that ASM forms first along bronchial branch stalks and only later at their tips, long implicating mesenchyme surrounding the stalk as the likely source of ASM progenitors (14, 31), as in the model above for VSM. However, an analysis of the expression pattern of an FGF10-regulated *lacZ* transgene suggested movement of tip mesenchyme along branch stalks (32), and the clone data and grafting experiments detailed below demonstrate that stalk mesenchyme does not contribute to ASM and that ASM progenitors are located exclusively at tips.

Labeled ASM cells in mesenchyme clones did not have siblings in the undifferentiated mesenchyme along the branch stalk, but rather sibling cells were located around and just ahead of the tip of the budding branch. These clones could be grouped into three categories based on the location and differentiation status of cells in the clones (Fig. 3, B to D). In the first category (18% of Rainbow clones analyzed at E13.5; clones 28, 32, 40 and 59 in Table 1), some cells in the clone, those furthest from the differentiated ASM cells, are undifferentiated mesenchyme at the branch tip several cell diameters away from the epithelium, whereas intervening cells in the clone are directly adjacent to the airway epithelium, elongating into the crescent shape characteristic of ASM cells but not expressing smooth muscle differentiation markers (Fig. 3B). In the second category (55% of E13.5 ASM-containing Rainbow clones; clones 29 to 31, 37, 41, 44, 46, 55, 58, 60,

64, and 65), some cells of the clone are elongated but marker negative as above, whereas other members of the clone are integrated into the ASM layer and express smooth muscle  $\alpha$ -actin (SMA) (Fig. 3C). In the third category (27% of E13.5 ASM-containing Rainbow clones; clones 14 to 19), all cells in the clone are definitive ASM cells expressing SMA (Fig. 3D). We propose that these categories represent three stages in a developmental series in which undifferentiated smooth muscle progenitors reside just ahead of the budding airway branch tip and are recruited to the ASM lineage. During recruitment by the airway epithelium, ASM progenitors adopt an elongated, crescent morphology and orient circumferentially around the bronchial branch and later turn on smooth muscle markers as the airway tip extends, leaving in its wake the differentiated contractile rings of smooth muscle cells along the airway stalk (Fig. 3F).



**Fig. 2. Cell dynamics of mesenchymal clones and structure of VSM clones.** **(A)** Scatter plots of each *Tbx4*<sup>LME-CreER</sup>;Rainbow mesenchyme clone (Table 1) showing percentage of differentiated cells (cells expressing ASM, VSM, mesothelial, or cartilage markers) in each clone as a function of clone size (number of cells in clone). Clones harvested at E13.5 (dots) have <200 cells and range from undifferentiated to fully differentiated, whereas clones harvested at E14.5 (squares) show the E13.5 pattern or are large (>200 cells) and mostly undifferentiated clones (dashed box), indicating the emergence of a highly proliferative fibroblast population soon after E13.5. Clone size is greater ( $P = 9 \times 10^{-7}$ ) and degree of differentiation less ( $P = 9 \times 10^{-5}$ ) for clones harvested at E14.5 compared with E13.5. **(B)** Violin plots of clones as above showing clone size (red dot, mean) as a function of cell type complexity (number of mesenchymal cell types). *n*, number of clones analyzed. **(C)** Ratio of cell number in E13.5 lungs of twinspace mesenchyme clones ( $n = 53$  clones; table S1) generated by HPRT-Cre;MADM, which differentially labels the two daughter cells of the original recombined cell with GFP or DsRed. Ratios were binned in increments of 0.5 (black bars) and shown with data from 10,000 simulated twinspace clones (gray bars; clone frequency values divided by 250) assuming a Poisson distribution in cell number. Note the overrepresentation of clones with strong asymmetries (ratios >5) relative to clones in Poisson simulation ( $P = 3 \times 10^{-101}$  for all ratios >5;  $P = 1 \times 10^{-18}$  with solo twinspace clones excluded). **(D and E)** Projections of confocal z stacks of twinspace clones 22 (D) and 47 (E) as above, immunostained for GFP (clone mark, green), DsRed (sibling cell clone mark,

magenta) and for E-cadherin (epithelium, blue), SMA (smooth muscle, red) and in (D) for PECAM (endothelium, also in blue). Values at bottom left show cell numbers in each twinspace; note the asymmetry (D) and outright loss of one twinspace (E). No pattern was observed between loss of a twinspace and the clone's location in the lung or position along a branch. **(F)** Mesenchyme clones generated as above showing how the arrangement of labeled cells in clones varies widely, from tight clusters (green-circled clone) to amorphous clouds of cells [yellow-circled clone, and clone in (D)]. A slight bias was noted for clones arranged along the length of a branch, as in the yellow-circled clone, but no correlation with specific lobes or branches formed by a particular mode was apparent. **(G and H)** Clones including VSM cells. In schematics (**G'** and **H'**), cells in each clone are numbered. Note the similar arrangement of cells, with undifferentiated mesenchyme cells (green in schematics) adjacent to SMA-expressing VSM cells (green with black dot in schematics). **(I)** Model of VSM recruitment. Progenitors (gray) in mesenchyme directly surrounding endothelial cells (blue) are recruited to form VSM of artery wall (red). Scale bars, 50  $\mu$ m [(C), (D), and (F)], 10  $\mu$ m [(G) and (H)].

The tip location of ASM progenitors was tested *ex vivo* by micrografting a tiny patch (~200 cells) of mesenchyme from a branch tip from *Zoanthus* sp. green fluorescent protein (ZsGreen)-expressing E11.5 and E12.5 donor lungs into tip positions in unlabeled littermate host lungs (Fig. 4A). Host lungs were cultured at 37°C for 4 days, and grafted cells were monitored by ZsGreen fluorescence. During the culture period, grafted tip cells moved proximally along the airway epithelium, with cells in 86% of cultures migrating 200  $\mu$ m or more (Fig. 4B and fig. S5, A to D). Many of the migrating cells near the airway epithelium differentiated into SMA-positive ASM cells (Fig. 4C). This confirms that ASM progenitors are located just ahead of branch tips and shows that they can migrate along the tip to reach the stalk and begin differentiating into ASM.

### A lineage boundary between stalk mesenchyme and ASM

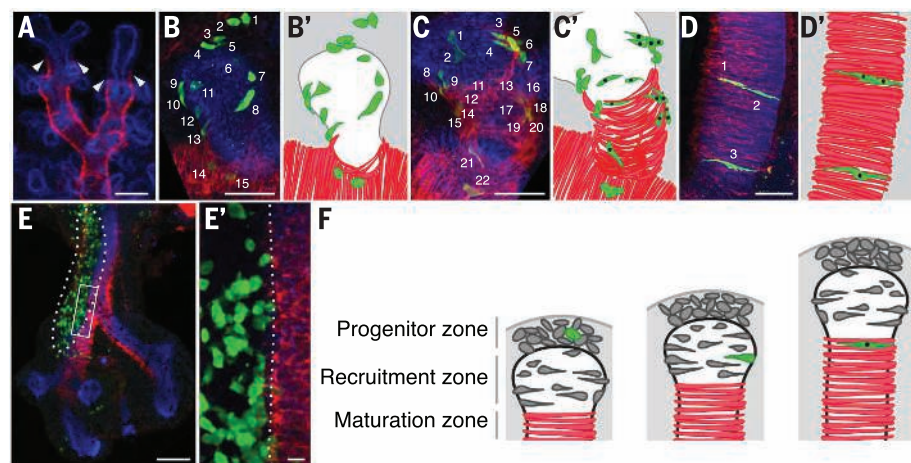
The failure of mesenchyme from airway flank to contribute to ASM is illustrated by large clones in which stalk mesenchyme was extensively labeled but mesenchyme at branch tips was not (Fig. 3E). Labeled cells in these clones surrounded and contacted the developing ASM, but they did not contribute to it; there was a lineage boundary at the edge of the ASM layer that flanking cells did not penetrate. Together with the above results,

these findings support a model in which (i) progenitors reside in a mobile progenitor pool near the distal tip of lung branches that moves with the growing branch tip; (ii) progenitor cells are progressively recruited from this pool onto the growing branch to form the ASM; and (iii) stalk mesenchyme cells are blocked from joining the ASM from the branch side.

### Reversal of the stalk lineage boundary to form new tip progenitor pools

The lung forms by three main modes of branching—domain branching, planar bifurcation, and orthogonal bifurcation (fig. S4) (33)—and each of these different types of branches is associated with a tip progenitor pool. As the branch number increases, so do the number of tip progenitor pools. In the case of branches formed by airway tip bifurcation, we infer that the tip progenitor pool simply splits with the bifurcation of the epithelial portion of the branch. However, late-forming domain branches emerge from the side of the parental branch after the smooth muscle layer along the parent branch has formed, so these branches must bud through an existing smooth muscle layer (34) (Fig. 4E). Rainbow clones generated with an ASM-specific CreER were found to be both multicellular and dispersed (fig. S6), showing that differentiated ASM cells of the parent branch remain

proliferative and migratory. To investigate whether the smooth muscle layer that forms along the new branches comes from recruitment of these differentiated ASM cells of the parent branch or from undifferentiated stalk mesenchyme, we used a smooth muscle-specific CreER (35) and a saturating dose of tamoxifen at E11.5 to label virtually all of the existing ASM cells of the parent branch just before the budding of new domain branches. We then allowed the new branches to grow out through the labeled smooth muscle and harvested the embryos 3 days later, after the smooth muscle layer had formed around the new branches. We found that the daughter branch smooth muscle was unlabeled (Fig. 4F), demonstrating that parental branch ASM is not a source of new smooth muscle cells along the daughter branch, nor does it substantially mix with daughter-branch ASM. Rather, the daughter branch induces a new tip progenitor pool from nonlineage labeled cells, presumably the stalk mesenchyme that flanks the ASM of the parent branch, overcoming the previous lineage restriction. Thus, with each new domain branching event a new source of ASM progenitors is induced, and a lineage boundary forms between the ASM layers of parental and daughter branches. The maintenance of this ASM lineage boundary also implies that any proliferation and cell movement within the formed ASM remains local and rarely crosses branch boundaries.



**Fig. 3. Clonal analysis reveals ASM progenitor niche, dynamics, and lineage boundaries.** (A) Airway branches of whole-mount E13.5 mouse lung immunostained for E-cadherin (epithelium, blue) and smooth muscle  $\alpha$ -actin (smooth muscle, red). SMA-positive ASM fibers cover airway stalks but are excluded from tips. Arrowheads, distal boundary of smooth muscle fibers. (B to D) ASM-containing clones in E13.5 lungs immunostained for tissue markers as in (A) and for GFP clone marker (green). Cells in clones are numbered. In schematics (B' to D'), clone cells are green, and those that are SMA-positive are marked with black dots. ASM clones fall into three classes: those with undifferentiated mesenchyme at the distal tip plus elongating but SMA-negative cells around the tip epithelium [(B) and (B')]; those with elongated but SMA-negative cells plus SMA-positive ASM cells [(C) and (C')]; and those with all cells SMA-positive [(D) and (D')]. Examples shown were chosen to illustrate the morphologies and distribution of clone cells; differences in cell number between clones are arbitrary. (E) A large clone (outlined) with extensive labeling of mesenchyme along the trachea and right primary airway stalk at E11.5. (E') Close-up of boxed region. Note the lack of recruitment into the ASM layer from labeled mesenchyme at the branch side, with the lineage boundary marked by a dotted line. (F) Three-step model of ASM formation. The fate of one progenitor (green) is highlighted. Proliferating progenitors are located in mesenchyme at the branch tip (progenitor zone), initiate differentiation through interaction with epithelium (recruitment zone), and mature along the airway stalk (maturation zone) as the branch extends. Scale bars, 100  $\mu$ m [(A) and (E)], 50  $\mu$ m [(B) to (D)], 10  $\mu$ m (E').

### A focal Wnt signal can prime stalk mesenchyme to form a tip progenitor pool

To determine whether all signals required to convert naïve mesenchyme into a tip progenitor pool emanate from the epithelial branch tip, we exposed stalk mesenchyme to epithelial branch tips by *ex vivo* grafting. Small patches of stalk mesenchyme from ZsGreen-expressing donor lungs were micrografted into tip positions of unlabeled littermate hosts (Fig. 4A), thereby exposing grafted cells to any signals coming from tip epithelium or mesenchyme flanking the graft site. During the 4-day culture period, the grafted stalk cells in nearly all grafts (89%;  $n = 27$  grafts) remained at the graft site with little or no movement along the airway (Fig. 4B) or differentiation to ASM (Fig. 4D). Thus, signals from tip epithelium and flanking mesenchyme are not sufficient to induce stalk mesenchyme to form a tip progenitor pool.

This failure of grafted stalk mesenchyme to respond to tip epithelial cues suggests that factors normally found in tip mesenchyme are required to induce or maintain the tip progenitor pool. To identify signals capable of replacing this required tip mesenchyme factor, we grafted small pieces of labeled stalk mesenchyme to branch tips as above, then implanted a ligand-soaked bead at the distal tip of the grafted tissue. Host lungs were placed into culture and the position of the engrafted cells was monitored for 4 days, after which samples were fixed and stained for differentiated cell types. Grafted stalk cells in nearly all grafts implanted with phosphate-buffered saline (PBS)-soaked control beads (88%;  $n = 8$  grafts) remained at the graft site, as expected (Fig. 5, A and D). However, grafted cells in grafts with beads



soaked in WNT1 (89%;  $n = 19$  grafts) migrated 100  $\mu\text{m}$  or more along the growing airway and contributed to the smooth muscle layer (Fig. 5, C and D, and fig. S5E), an effect that was partially blocked by inclusion in the culture medium of extracellular Wnt antagonists Dkk1 or Sfrp1 (Fig. 5D). Indeed, stalk mesenchyme grafts exposed to WNT1 beads were indistinguishable from grafts of tip mesenchyme (Fig. 5, C and D, and Fig. 4C), demonstrating that a distal Wnt signal is sufficient to overcome the barrier to migration and smooth muscle recruitment observed in untreated stalk mesenchyme grafts. This effect was not seen with FGF10-soaked beads ( $n = 8$  grafts) (Fig. 5, B and D), although FGF10 is a factor prominently expressed in tip mesenchyme and it has been implicated in smooth muscle recruitment (32, 36).

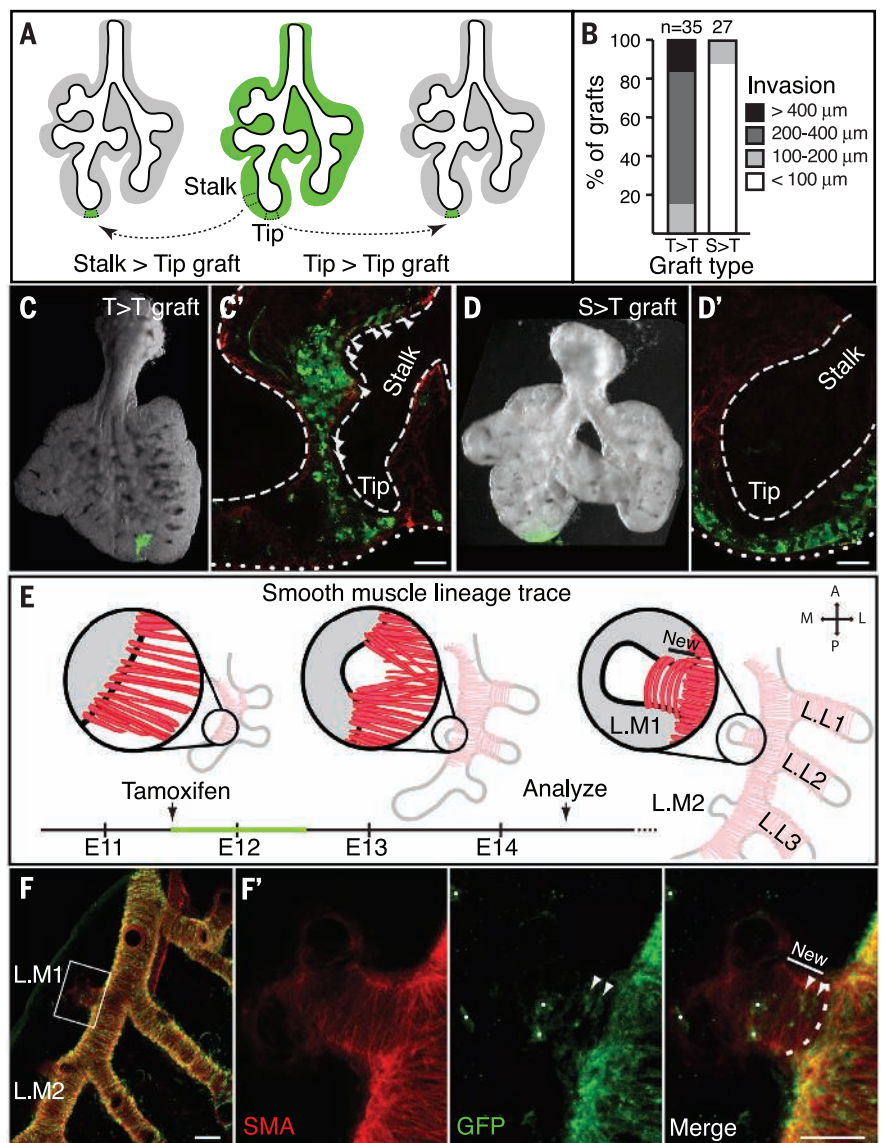
To examine the molecular effects of Wnt signaling on stalk mesenchyme, we used oligonu-

cleotide microarrays to transcriptionally profile stalk mesenchyme cultured with or without WNT1 and identified a Wnt response signature of 306 genes whose expression changed by a factor of two or more (table S3). The Wnt response signature was enriched for WNT signaling genes as expected, and also for cell motility and migration, cell projection, and contractile fiber genes (table S4), which presumably mediate the observed migratory and differentiative effects of WNT1. Comparison of the expression profiles of untreated stalk and tip mesenchyme showed that Wnt response signature genes were expressed at higher levels in tip mesenchyme (fig. S7), supporting a role for Wnt signaling in defining tip mesenchyme potential in vivo.

Because a Wnt ligand was sufficient to transform grafted stalk mesenchyme into cells whose behavior is indistinguishable from that of the tip progenitor pool, and because higher expression

of Wnt response signature genes distinguishes tip mesenchyme, we tested the effect of increased levels of canonical Wnt signaling in the mesenchyme along branch stalks using  $\text{Tbx4}^{\text{LME}}\text{Cre}$  to express a constitutively active form of  $\beta$ -catenin in the mesenchyme. In such lungs, the distal vascular plexus, which in control lungs of the same stage is a simple endothelial net, became wrapped with ectopic smooth muscle cells expressing the VSM marker NG2 (Fig. 5, F to H), an arrangement reminiscent of the pericytes surrounding mature capillaries. This precocious recruitment of VSM to the distal plexus is accompanied by more subtle patterning changes in the smooth muscle layer around airways: The normally strictly parallel fibers oriented orthogonal to the branch axis (Fig. 5E) were disorganized in the mutant (Fig. 5F), and the cells were less tightly apposed, resulting in a thicker, less-condensed ASM layer.

**Fig. 4. Tissue grafts and lineage tracing show that ASM progenitors are located in tip mesenchyme and form anew with new domain branches.** (A) Graft strategy. Mesenchyme excised from the stalk or the tip of ZsGreen-expressing donor E11.5 lung were grafted into the tip of an unlabeled littermate host lung, and grafted cells (green) were traced in culture. (B) Quantification of progenitor cell migration from the graft in tip-to-tip (T>T) or stalk-to-tip (S>T) mesenchyme grafts, measuring the distance from the graft margin to the furthest invading graft cell, along the shortest inferred migration path.  $n$ , number of grafts. Note the robust invasion of tip-derived cells (all grafts showed >100  $\mu\text{m}$  migration) relative to stalk-derived cells (11% showed >100  $\mu\text{m}$  migration). (C and C') Stereomicroscope image showing recipient lung 4 days after tip-to-tip micrograft (C), and single confocal optical slice of grafted region (C'). Note in (C') that tip cells have migrated away from lung margin (dotted line) along the airways (dashed lines) and that some graft cells (arrowheads) are expressing ASM differentiation marker smooth muscle actin (red). (D) S>T mesenchyme graft. Grafted cells fail to leave the graft site and do not migrate along airways or express smooth muscle actin (D'). (E) Strategy of lineage trace of smooth muscle of parent branch of L.M1, a late-forming domain branch. Budding L.M1 (highlighted in insets) is the first medial domain branch formed along the left primary airway; it buds through the existing ASM layer of the parent branch and forms its own ASM layer. Saturating tamoxifen (3 mg) was administered at E11.5 to SM-MHC (smooth muscle myosin heavy chain)-CreER;mTmG to label differentiated ASM of the parent branch and early-forming lateral branches (L.L1 to L.L3) with heritable GFP expression (green bar on timeline, with tamoxifen assumed active for 24 hours) before emergence of an L.M1 bud at E12.5. Lungs were analyzed at E14.5 after L.M1 has budded and formed its ASM ("new" ASM, black bar in inset). (F) Confocal z-stack projections (ventral view) showing close-up of boxed region (F') of an SM-MHC-CreER;mTmG E14.5 embryo treated as above and immunostained for lineage trace of smooth muscle of the parent branch (GFP, green) and all ASM (SMA, red). Note that the ASM of the L.M1 daughter branch is not lineage labeled [border highlighted with dashed line in (F') merge], indicating a new progenitor pool formed ahead of the new branch that is lineally independent of existing ASM. There is no cell mixing across the smooth muscle lineage boundary (dashed line), with the exception of two lineage-labeled cells [arrowheads in (F')] in ASM of L.M1. Similar results are seen for branch L.M2. Dots, lineage-labeled non-ASM cells. Scale bars, 50  $\mu\text{m}$ .



In addition, smooth muscle cells were observed connecting the ASM layer with the ectopic VSM layer wrapping the vascular plexus (Fig. 5G), a feature never observed in control lungs at any stage.

## Discussion

Our lung mesenchyme-specific clonal analysis and micrografting experiments show that ASM progenitors are not broadly distributed but rather are confined to dynamic local pools around branch tips that move and bifurcate with the tips and from which progenitors are continuously recruited onto airway stalks, where they differentiate into smooth muscle. Clonal analysis also revealed a lineage boundary that prevents mesenchymal cells surrounding the stalks from becoming ASM from branch sides. However, new progenitor pools are induced in these regions around late-forming domain branches, creating a smooth muscle progenitor pool for each new branch whose progeny do not intermix with those of the parent branch. These progenitor pools require a budding airway tip and a priming signal that can be provided by a focal Wnt source, and delocalized Wnt pathway activity *in vivo* expands or alters the progenitor pool, causing ectopic smooth muscle formation on nearby plexus endothelial cells. Thus, ASM progenitors are organized into localized pools that are carefully regulated by surrounding signals.

These localized mesenchymal progenitor pools share several features with classical epithelial stem cell niches, such as those that maintain adult hair follicles and gut epithelium. Each is a focal

collection of progenitors that resides in a specialized signaling environment that controls their proliferation as well as their recruitment out of the niche and differentiation into specialized cells (11, 37), thereby coordinating the generation of a critical cell type with organ growth or maintenance. But unlike classical epithelial niches, which are typically located at fixed positions with stem cells segregated from other cell types (38), these mesenchymal progenitor pools are highly dynamic and not bounded in any obvious way. A new smooth muscle progenitor pool is established with each new budding branch, and it moves and splits as the tip extends and bifurcates, with rapidly cycling and highly motile progenitors leaving the tip pool after only short and variable stays. Whereas epithelial stem cells self-renew (39, 40), the smooth muscle progenitor pools are gradually depleted during embryogenesis, although some might remain at the end of branching to serve as adult stem cells in airway wall maintenance and remodeling.

Progenitors in lung mesenchyme are not all organized into such focal progenitor niches (Fig. 6). Mesothelial progenitors populate the full mesenchymal surface, and clones remained coherent and restricted to the surface and the mesothelial lineage early in lung development (before E13.5); later, this barrier is lifted as dispersed clones containing mesothelium and other mesenchymal derivatives were recovered. VSM progenitors are even more broadly distributed, matching closest the conventional view of delocalized mesenchymal progenitors. However, even in this case, there is

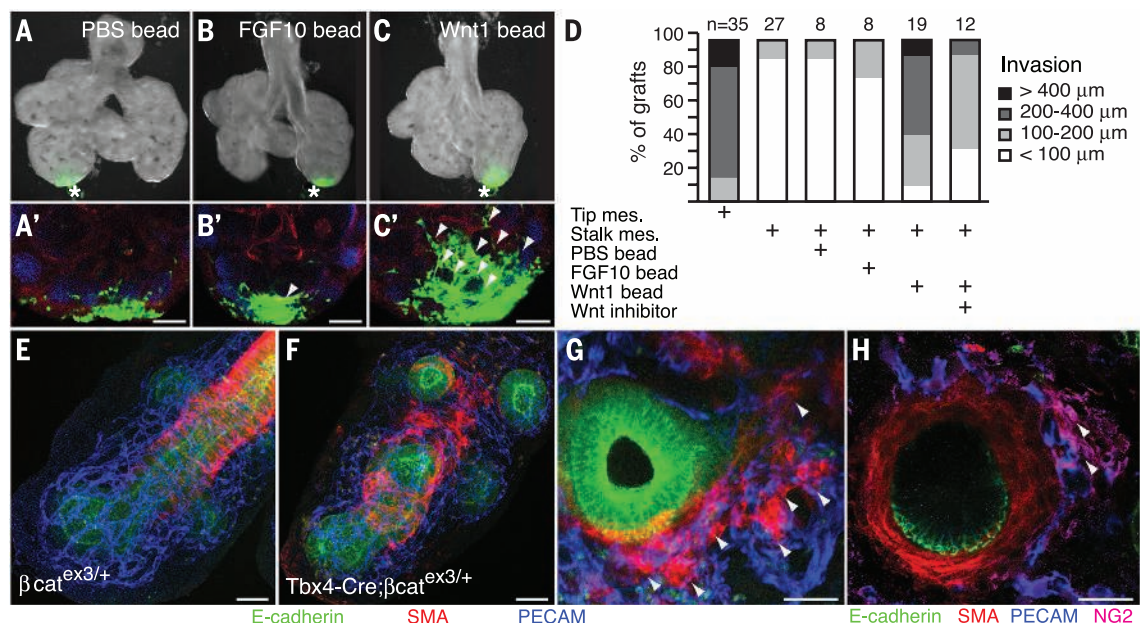
precise temporal and spatial control of progenitor recruitment, as VSM begins to form only at positions where the early vascular plexus has been replaced by the larger and more regular arterial and venous endothelial tubes. The mixed clones containing multiple differentiated cell types (clones 54 to 78, Table 1) demonstrate that daughter cells of early progenitors can disperse to seed multiple progenitor pools. Thus, lung mesenchyme is neither a large homogeneous progenitor pool, as it has classically been viewed, nor a collection of discrete, isolated, and unchanging progenitor niches. Rather, each differentiated cell type has a distinct mode of progenitor recruitment and regulated boundaries, which are modulated temporally and spatially, such that delocalization of the control signals disrupts the progenitor pools and therefore the patterning of stromal and support tissues. It will be important to elucidate the full set of signals that controls individual cells in each of these pools, including the specific Wnt(s) and other signals that control ASM progenitors, and how these signals change during development and in adult life. It will also be valuable to use the clonal and other approaches described here to explore the full diversity and dynamics of mesenchyme progenitor pools in other organs and disease contexts.

## Materials and methods

### Mice

Mouse lines used and details of the cloning of the Tbx4<sup>LME</sup>-Cre and -CreERT2 transgenes and construction of the transgenic lines are described

**Fig. 5. Effect of Wnt bead implants and transgenic activation of Wnt pathway show that Wnt signaling primes ASM progenitors.** (A to C) Stereo-scope images of S>T mesenchyme grafts performed as in Fig. 4D, except also implanted with control (A) or ligand-soaked [(B) and (C)] agarose beads as indicated. Asterisk, position of bead. (A' to C') Single confocal slices of the graft region after immunostaining for GFP (graft marker, green), E-cadherin (epithelium, blue), and SMA (ASM, red). Graft cells exposed to PBS- or FGF10-



soaked beads showed little or no migration along airways [(A') and (B')], whereas those exposed to WNT1-soaked beads (C') migrated along airways where some (arrowheads) differentiated to ASM (separate channels are shown in Fig. S5E). (D) Quantification of graft cell invasion in untreated grafts (reproduced from Fig. 4B) and after exposure to ligand-soaked beads and Wnt inhibitors, as indicated. WNT1 beads (column 5) induce stalk mesenchyme invasion at levels comparable to that seen in T>T grafts (column 1), and the effect is reduced in the presence of 100 ng/ml Wnt inhibitors DKK1 or SFRP1 (column 6). (E to H) Right middle lobes of E12.5 control (E) and Tbx4<sup>LME</sup>-

Cre;βcat<sup>ex3/+</sup> lungs that express stabilized β-catenin to activate the canonical Wnt pathway throughout the mesenchyme [(F) to (H)] immunostained for E-cadherin (epithelium, green), SMA (smooth muscle, red), and PECAM (endothelium, blue). Wnt pathway activation results in disorganized smooth muscle around airways (F) and ectopic wrapping of vascular plexus by cells continuous with ASM [arrowheads in (G)]. (H) Single confocal slice of branch as in (G), showing that the ectopic smooth muscle wrapping the plexus (arrowhead) expresses VSM marker NG2 (magenta). Scale bars, 50 μm [(A) to (E)], 25 μm [(F) and (G)].



in the supplementary materials. All animal experiments were performed in accordance with national and institutional guidelines.

### Clonal analysis

Clones were visualized using two independent multicolor cell labeling systems, MADM (27) and Rainbow (25). In MADM, Cre-mediated interchromosomal recombination between MADM-RG and MADM-GR chromosomes in HPRT-Cre;MADM-GR/MADM-RG embryos generates MADM-GG (full-length GFP) and MADM-RR (full-length myc-epitope-tagged DsRed) chromosomes, which can either segregate to the same daughter cell to give yellow clones or to separate sibling daughter cells to form twin spots with green GFP daughter cells and red myc-epitope-tagged DsRed daughter cells. Recombination is extremely rare and not temporally controlled; however, no labeled cells were detected in lungs before E11.5. Lungs were harvested between E11.5 and E13.5, the latest stage at which individual clones could be resolved. Whole-mount lungs were fixed and stained with antisera specific for clone markers GFP and myc-epitope-tagged DsRed plus a panel of tissue-specific markers, as described in the supplementary materials. Whole-mount immunostained lungs were imaged by confocal fluorescence microscopy. When the recombination of MADM chromosomes was driven with  $Tbx4^{LME}$ -Cre, labeled cells were not observed until after the developmental window considered here.

In the Rainbow system (25), pregnant  $Tbx4^{LME}$ -CreER;Rosa26R-Rainbow females were given a single intraperitoneal injection of 0.5 to 0.7 mg tamoxifen (Sigma) dissolved in corn oil (Sigma) at the indicated gestational ages, and embryos were harvested several days later, as indicated. The low doses of tamoxifen induced rare recombination events; most lungs contained no labeled cells ( $n = 177$  of 351 lungs assayed). Embryos were dissected, and their lungs were fixed, cryosectioned, and stained with SMA directly conjugated to Cy5 as described below and examined using fluorescent filters specific to each fluorophore. Only lungs containing one, or at most two, of the potential three-colored recombination products were used for further analysis ( $n = 56$ ). The number, location, and differentiation state of the cells in each clone were recorded, and selected clones ( $n = 70$ ) were analyzed by confocal fluorescence microscopy.

### Grafting, bead implantation, and organ culture

Small blocks of mesenchyme ( $<0.5$  mm) containing roughly  $200 \pm 50$  cells were dissected from ZsGreen-expressing donor lungs at the indicated positions and micrografted to branch tips of wild-type littermate host lungs at E11.5 to E12.5. Grafted lungs were placed in organ culture in DMEM/F12 with 10% fetal bovine serum for 3 to 4 days at  $37^{\circ}\text{C}$ , and the invasion and differentiation of grafted cells were assessed without fixation by flu-

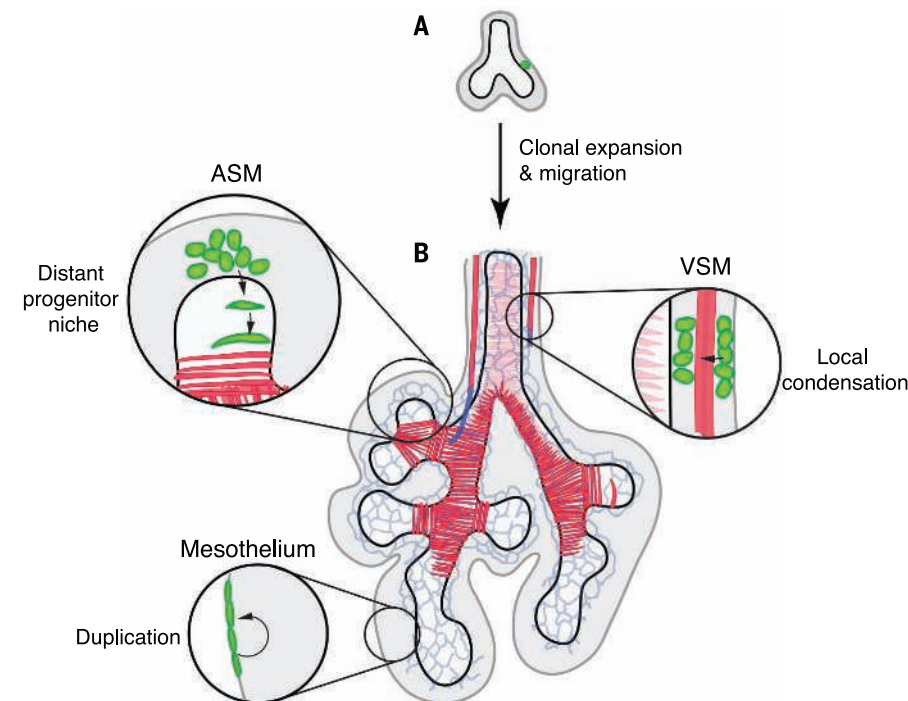
orescence stereomicroscopy and then by confocal fluorescence microscopy after fixation and immunohistochemistry. For bead implantation, after healing of the graft site overnight in culture, an agarose bead soaked in the indicated ligand was implanted at the graft site. For inhibitor treatments, 1 hour after bead implantation, organ cultures were transferred to growth medium containing the indicated concentrations of inhibitor.

### Transcriptional profiling and statistics

Mesenchyme blocks from tip and stalk were dissected as above, and total RNA was isolated, amplified, labeled, and used to probe Mouse Genome 430 2.0 Arrays (Affymetrix). To define the Wnt response signature, stalk mesenchyme pieces dissected as above were cultured overnight in either control media or media containing 400 ng/ml WNT1 before RNA isolation and analysis, and genes with expression differences of twofold or greater between the conditions were identified using Affymetrix software. Details of the experimental and statistical analysis are given in the supplementary materials.

### REFERENCES AND NOTES

1. K. MacCord, Mesenchyme. *Embryo Project Encyclopedia*, <http://embryo.asu.edu/pages/mesenchyme> (2012).
2. C. W. Prentiss, *A Laboratory Manual and Text-book of Embryology* (W. B. Saunders Company, Philadelphia, London, 1915).
3. L. B. Arey, *Developmental Anatomy, a Textbook and Laboratory Manual of Embryology* (W. B. Saunders Company, Philadelphia, London, ed. 5, 1946).
4. C. Grobstein, Inductive epitheliomesenchymal interaction in cultured organ rudiments of the mouse. *Science* **118**, 52–55 (1953). doi: [10.1126/science.118.3054.52](https://doi.org/10.1126/science.118.3054.52); pmid: [13076182](https://pubmed.ncbi.nlm.nih.gov/13076182/)
5. B. S. Spooner, N. K. Wessells, Mammalian lung development: Interactions in primordium formation and bronchial morphogenesis. *J. Exp. Zool.* **175**, 445–454 (1970). doi: [10.1002/jez.1401750404](https://doi.org/10.1002/jez.1401750404); pmid: [5501462](https://pubmed.ncbi.nlm.nih.gov/5501462/)
6. N. A. Bhowmick, E. G. Neilson, H. L. Moses, Stromal fibroblasts in cancer initiation and progression. *Nature* **432**, 332–337 (2004). doi: [10.1038/nature03096](https://doi.org/10.1038/nature03096); pmid: [15549095](https://pubmed.ncbi.nlm.nih.gov/15549095/)
7. J. B. Sneddon, M. Borowiak, D. A. Melton, Self-renewal of embryonic-stem-cell-derived progenitors by organ-matched mesenchyme. *Nature* **491**, 765–768 (2012). pmid: [23041930](https://pubmed.ncbi.nlm.nih.gov/23041930/)
8. J. R. Rock *et al.*, Multiple stromal populations contribute to pulmonary fibrosis without evidence for epithelial to mesenchymal transition. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E1475–E1483 (2011). doi: [10.1073/pnas.1117988108](https://doi.org/10.1073/pnas.1117988108); pmid: [22123957](https://pubmed.ncbi.nlm.nih.gov/22123957/)
9. R. Berair, R. Saunders, C. E. Brightling, Origins of increased airway smooth muscle mass in asthma. *BMC Med.* **11**, 145 (2013). doi: [10.1186/1741-7015-11-145](https://doi.org/10.1186/1741-7015-11-145); pmid: [23742314](https://pubmed.ncbi.nlm.nih.gov/23742314/)
10. C. Blanpain, E. Fuchs, Epidermal homeostasis: A balancing act of stem cells in the skin. *Nat. Rev. Mol. Cell Biol.* **10**, 207–217 (2009). doi: [10.1038/nrm2636](https://doi.org/10.1038/nrm2636); pmid: [19209183](https://pubmed.ncbi.nlm.nih.gov/19209183/)
11. H. Clevers, The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274–284 (2013). doi: [10.1016/j.cell.2013.07.004](https://doi.org/10.1016/j.cell.2013.07.004); pmid: [23870119](https://pubmed.ncbi.nlm.nih.gov/23870119/)
12. B. D. Humphreys *et al.*, Fate tracing reveals the pericyte and not epithelial origin of myofibroblasts in kidney fibrosis. *Am. J. Pathol.* **176**, 85–97 (2010). doi: [10.2353/ajpath.2010.090517](https://doi.org/10.2353/ajpath.2010.090517); pmid: [20008127](https://pubmed.ncbi.nlm.nih.gov/20008127/)
13. J. V. Taderera, Control of lung differentiation in vitro. *Dev. Biol.* **16**, 489–512 (1967). doi: [10.1016/0012-1606\(67\)90061-9](https://doi.org/10.1016/0012-1606(67)90061-9); pmid: [6053290](https://pubmed.ncbi.nlm.nih.gov/6053290/)
14. M. Weaver, L. Batts, B. L. Hogan, Tissue interactions pattern the mesenchyme of the embryonic mouse lung. *Dev. Biol.* **258**, 169–184 (2003). doi: [10.1016/S0012-1606\(03\)00117-9](https://doi.org/10.1016/S0012-1606(03)00117-9); pmid: [12781691](https://pubmed.ncbi.nlm.nih.gov/12781691/)
15. J. Mao, B. M. Kim, M. Rajurkar, R. A. Shivdasani, A. P. McMahon, Hedgehog signaling controls mesenchymal growth in the developing mammalian digestive tract. *Development* **137**, 1721–1729 (2010). doi: [10.1242/dev.044586](https://doi.org/10.1242/dev.044586); pmid: [20430747](https://pubmed.ncbi.nlm.nih.gov/20430747/)



**Fig. 6. Diverse mechanisms generate mesenchymal cell derivatives.** (A) Single labeled mesenchyme cells (green) proliferate (clonal expansion), and daughter cells disperse to seed progenitor niches that generate stromal and support cells by different mechanisms. (B) ASM progenitors are recruited to branch stalks from a niche distal to the branch tip. VSM progenitors are recruited by local condensation of mesenchyme immediately surrounding the vessel. Mesothelium forms by self-duplication and limited spreading of progenitors. Black, airway epithelium; gray, mesenchyme; dark gray, mesothelium; red, smooth muscle; blue, endothelium.

16. A. M. Goss *et al.*, Wnt2 signaling is necessary and sufficient to activate the airway smooth muscle program in the lung by regulating myocardin/Mrtf-B and Fgf10 expression. *Dev. Biol.* **356**, 541–552 (2011). doi: [10.1016/j.ydbio.2011.06.011](https://doi.org/10.1016/j.ydbio.2011.06.011); pmid: [21704027](https://pubmed.ncbi.nlm.nih.gov/21704027/)
17. H. J. Snippert *et al.*, Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010). doi: [10.1016/j.cell.2010.09.016](https://doi.org/10.1016/j.cell.2010.09.016); pmid: [20887898](https://pubmed.ncbi.nlm.nih.gov/20887898/)
18. G. Mascré *et al.*, Distinct contribution of stem and progenitor cells to epidermal maintenance. *Nature* **489**, 257–262 (2012). doi: [10.1038/nature11393](https://doi.org/10.1038/nature11393); pmid: [22940863](https://pubmed.ncbi.nlm.nih.gov/22940863/)
19. T. J. Desai, D. G. Brownfield, M. A. Krasnow, Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* **507**, 190–194 (2014). doi: [10.1038/nature12930](https://doi.org/10.1038/nature12930); pmid: [24499815](https://pubmed.ncbi.nlm.nih.gov/24499815/)
20. E. E. Morrisey, B. L. Hogan, Preparing for the first breath: Genetic and cellular mechanisms in lung development. *Dev. Cell* **18**, 8–23 (2010). doi: [10.1016/j.devcel.2009.12.010](https://doi.org/10.1016/j.devcel.2009.12.010); pmid: [20152174](https://pubmed.ncbi.nlm.nih.gov/20152174/)
21. H. Zong, J. S. Espinosa, H. H. Su, M. D. Muzumdar, L. Luo, Mosaic analysis with double markers in mice. *Cell* **121**, 479–492 (2005). doi: [10.1016/j.cell.2005.02.012](https://doi.org/10.1016/j.cell.2005.02.012); pmid: [15882628](https://pubmed.ncbi.nlm.nih.gov/15882628/)
22. S. H. Tang, F. J. Silva, W. M. Tsark, J. R. Mann, A Cre/loxP-deleter transgenic line in mouse strain 129S1/SvJm. *Genesis* **32**, 199–202 (2002). doi: [10.1002/gene.10030](https://doi.org/10.1002/gene.10030); pmid: [11892008](https://pubmed.ncbi.nlm.nih.gov/11892008/)
23. D. B. Menke, C. Guenther, D. M. Kingsley, Dual hindlimb control elements in the Tbx4 gene and region-specific control of bone size in vertebrate limbs. *Development* **135**, 2543–2553 (2008). doi: [10.1242/dev.017384](https://doi.org/10.1242/dev.017384); pmid: [18579682](https://pubmed.ncbi.nlm.nih.gov/18579682/)
24. P. Bogard, thesis, Stanford University, Stanford, CA (2012).
25. Y. Rinkevich, P. Lindau, H. Ueno, M. T. Longaker, I. L. Weissman, Germ-layer and lineage-restricted stem/progenitors regenerate the mouse digit tip. *Nature* **476**, 409–413 (2011). doi: [10.1038/nature10346](https://doi.org/10.1038/nature10346); pmid: [21866153](https://pubmed.ncbi.nlm.nih.gov/21866153/)
26. F. H. Crick, P. A. Lawrence, Compartments and polyclones in insect development. *Science* **189**, 340–347 (1975). doi: [10.1126/science.806966](https://doi.org/10.1126/science.806966); pmid: [806966](https://pubmed.ncbi.nlm.nih.gov/806966/)
27. E. Cano, R. Carmona, R. Muñoz-Chápuli, Wt1-expressing progenitors contribute to multiple tissues in the developing lung. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **305**, L322–L332 (2013). doi: [10.1152/ajplung.00424.2012](https://doi.org/10.1152/ajplung.00424.2012); pmid: [23812634](https://pubmed.ncbi.nlm.nih.gov/23812634/)
28. R. Dixit, X. Ai, A. Fine, Derivation of lung mesenchymal lineages from the fetal mesothelium requires hedgehog signaling for mesothelial cell entry. *Development* **140**, 4398–4406 (2013). doi: [10.1242/dev.098079](https://doi.org/10.1242/dev.098079); pmid: [24130328](https://pubmed.ncbi.nlm.nih.gov/24130328/)
29. R. V. Pearse 2nd, P. J. Scherz, J. K. Campbell, C. J. Tabin, A cellular lineage analysis of the chick limb bud. *Dev. Biol.* **310**, 388–400 (2007). doi: [10.1016/j.ydbio.2007.08.002](https://doi.org/10.1016/j.ydbio.2007.08.002); pmid: [17888899](https://pubmed.ncbi.nlm.nih.gov/17888899/)
30. D. M. Greif *et al.*, Radial construction of an arterial wall. *Dev. Cell* **23**, 482–493 (2012). doi: [10.1016/j.devcel.2012.07.009](https://doi.org/10.1016/j.devcel.2012.07.009); pmid: [22975322](https://pubmed.ncbi.nlm.nih.gov/22975322/)
31. J. M. Flint, The development of the lungs. *Am. J. Anat.* **6**, 1–137 (1906). doi: [10.1002/aja.1000060102](https://doi.org/10.1002/aja.1000060102)
32. A. A. Mailleux *et al.*, Fgf10 expression identifies parabronchial smooth muscle cell progenitors and is required for their entry into the smooth muscle cell lineage. *Development* **132**, 2157–2166 (2005). doi: [10.1242/dev.01795](https://doi.org/10.1242/dev.01795); pmid: [15800000](https://pubmed.ncbi.nlm.nih.gov/15800000/)
33. R. J. Metzger, O. D. Klein, G. R. Martin, M. A. Krasnow, The branching programme of mouse lung development. *Nature* **453**, 745–750 (2008). doi: [10.1038/nature07005](https://doi.org/10.1038/nature07005); pmid: [18463632](https://pubmed.ncbi.nlm.nih.gov/18463632/)
34. R. Metzger, thesis, Stanford University, Stanford, CA (2007).
35. A. Wirth *et al.*, G12-G13-LARG-mediated signaling in vascular smooth muscle is required for salt-induced hypertension. *Nat. Med.* **14**, 64–68 (2008). doi: [10.1038/nm1666](https://doi.org/10.1038/nm1666); pmid: [18084302](https://pubmed.ncbi.nlm.nih.gov/18084302/)
36. S. P. De Langhe, G. Carraro, D. Warburton, M. K. Hajihosseini, S. Bellusci, Levels of mesenchymal FGFR2 signaling modulate smooth muscle progenitor cell commitment in the lung. *Dev. Biol.* **299**, 52–62 (2006). doi: [10.1016/j.ydbio.2006.07.001](https://doi.org/10.1016/j.ydbio.2006.07.001); pmid: [16989802](https://pubmed.ncbi.nlm.nih.gov/16989802/)
37. T. Tumber *et al.*, Defining the epithelial stem cell niche in skin. *Science* **303**, 359–363 (2004). doi: [10.1126/science.1092436](https://doi.org/10.1126/science.1092436); pmid: [14671312](https://pubmed.ncbi.nlm.nih.gov/14671312/)
38. T. Sato *et al.*, Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469**, 415–418 (2011). doi: [10.1038/nature09637](https://doi.org/10.1038/nature09637); pmid: [21113151](https://pubmed.ncbi.nlm.nih.gov/21113151/)
39. C. Blanpain, W. E. Lowry, A. Geoghegan, L. Polak, E. Fuchs, Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell* **118**, 635–648 (2004). doi: [10.1016/j.cell.2004.08.012](https://doi.org/10.1016/j.cell.2004.08.012); pmid: [15339667](https://pubmed.ncbi.nlm.nih.gov/15339667/)
40. N. Barker *et al.*, Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003–1007 (2007). doi: [10.1038/nature06196](https://doi.org/10.1038/nature06196); pmid: [17934449](https://pubmed.ncbi.nlm.nih.gov/17934449/)

## ACKNOWLEDGMENTS

This work was supported by National Research Service Award (NRSA) postdoctoral fellowship F32HL083645 (M.E.K.), NIH grants R01HL075769 and U01HL099995, and the Vera M. Wall Center for Pulmonary Vascular Disease at Stanford University (M.A.K.); Damon Runyon and Parker B. Francis postdoctoral fellowships (F.H.E.); NRSA postdoctoral fellowship F32HD048006 (D.B.M.); and NIH Center of Excellence in Genomic Studies grant 5P50HG2568 (D.M.K.). M.A.K. and D.M.K. are investigators of the Howard Hughes Medical Institute. M.E.K. and M.A.K. designed the experiments, performed data analysis, and wrote the manuscript. M.E.K. performed the experiments. M.E.K. and P.E.B. made the Tbx4<sup>LME</sup> transgenic mice. F.H.E. analyzed gene expression at branch tips. D.B.M. and D.M.K. isolated the Tbx4 enhancer element. The data reported in this paper are tabulated in the supplementary materials, and expression profiling data sets have been deposited to the Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo)) (GSE62039). The authors thank D. Riordan for help with statistical analysis and L. Luo and I. Weissman for generously sharing mouse lines before publication.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/346/6211/1258810/suppl/DC1](http://www.sciencemag.org/content/346/6211/1258810/suppl/DC1)  
Supplementary Text  
Figs. S1 to S7  
Tables S1 to S5  
References (41–51)

16 July 2014; accepted 3 October 2014  
10.1126/science.1258810



## REPORTS

## NANOPHOTONICS

# Plasmoelectric potentials in metal nanostructures

Matthew T. Sheldon,<sup>1</sup> Jorik van de Groep,<sup>2</sup> Ana M. Brown,<sup>1</sup> Albert Polman,<sup>2</sup> Harry A. Atwater<sup>1\*</sup>

The conversion of optical power to an electric potential is of general interest for energy applications and is typically obtained via optical excitation of semiconductor materials. We developed a method for achieving electric potential that uses an all-metal geometry based on the plasmon resonance in metal nanostructures. In arrays of gold nanoparticles on an indium tin oxide substrate and arrays of 100-nanometer-diameter holes in 20-nanometer-thick gold films on a glass substrate, we detected negative and positive surface potentials during monochromatic irradiation at wavelengths below or above the plasmon resonance, respectively. We observed plasmoelectric surface potentials as large as 100 millivolts under illumination of 100 milliwatts per square centimeter. Plasmoelectric devices may enable the development of all-metal optoelectronic devices that can convert light into electrical energy.

Noble metal nanostructures display remarkable optical properties that arise from the coupling of incident light to the collective motion of the conduction electrons. The excitation, propagation, and localization of these plasmons can be tailored by nanoscale control of metal size and shape (1). Metal nanostructures exhibiting subwavelength optical confinement (2) have enabled nanoscale photonic waveguides, modulators, enhancement of second-harmonic generation, light-trapping structures for photovoltaics, and biological labeling techniques (1, 2). Coupling of laser light to plasmonic structures can also result in efficient localized heating, because the resonant absorption cross section for plasmonic nanostructures is much larger than their physical cross section.

Recent work has demonstrated control over the plasmon resonance frequency ( $\omega_p$ ) of metal nanostructures when an external electrostatic field alters the carrier density in the metal (3). Increasing the carrier density in a noble metal nanoparticle leads to a blueshift of the resonance, whereas decreasing it leads to a redshift, as illustrated for a 20-nm-diameter Ag nanoparticle in vacuum (Fig. 1). The reverse effect—the optical generation of an electrostatic potential due to an optically driven change in carrier density in a plasmonic nanostructure—has so far not been observed. Thermodynamically, however, such a plasmoelectric effect is expected to occur. We obtained direct experimental evidence of plasmoelectric potentials in the range 10 to 100 mV on colloidal assemblies and plasmonic light-

harvesting device geometries, in qualitative agreement with a thermodynamic model. These results may provide a new route to convert optical energy into electrical power.

Figure 2A shows the measured scattering spectrum for 60-nm Au colloids on indium tin oxide (ITO)/glass, showing a clear plasmon resonance around  $\lambda = 550$  nm. To probe the local static potential difference between the tip and sample surface at room temperature, we used Kelvin probe force microscopy (KPFM) (4) with a conductive atomic force microscope tip in noncontact mode (Fig. 2). While the illumination wavelength was gradually scanned through the plasmon resonance spectrum, from 480 to 650 nm ( $\sim 1$  nm  $s^{-1}$ ), we probed the potential on the illuminated Au nanoparticle array (5). A clear optically induced surface potential was observed, which varied with illumination wavelength (Fig. 2B). We observed negative induced potentials during excitation to the blue side of the neutral-particle plasmon resonance wavelength near 550 nm, and positive potentials during excitation on the red side of the resonance, with the measured potential changing sign near the peak of the plasmon resonance.

To model the experimentally observed plasmoelectric effect, we consider a metal nanostructure placed on a grounded conducting substrate that is illuminated with monochromatic radiation at a wavelength just below the plasmon resonance  $\lambda_p$ . Random charge fluctuations between particle and substrate will cause the plasmon resonance spectrum to vary by minute amounts (Fig. 1). If an electron is randomly added to the particle, the resonance will shift toward the blue, leading to a concomitant increase in light absorption of the particle, which in turn leads to a small increase in the nanoparticle temperature. The changes in number of electrons  $N$  and temperature  $T(N)$  change the free energy  $F$  of the particle, and an

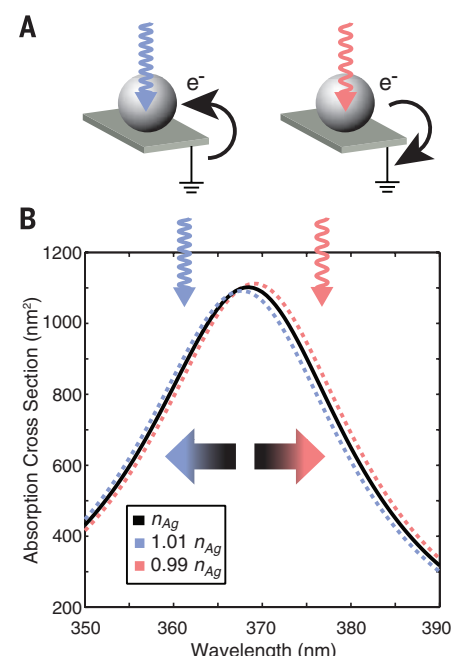
equilibrium charge density is achieved when the free energy is at a minimum:

$$\frac{\partial F(N, T)}{\partial N} = \left( \frac{\partial F}{\partial N} \right)_T + \left( \frac{\partial F}{\partial T} \right)_N \frac{dT}{dN} = 0 \quad (1)$$

Here, we assume that both the intensity and wavelength of the illumination are constant. Using, by definition,  $(\partial F/\partial N)_T \equiv \mu$ , where  $\mu$  is the electrochemical potential, and  $-(\partial F/\partial T)_N \equiv S$ , where  $S$  is the entropy of the particle, we find

$$\mu = S \frac{dT}{dN} \quad (2)$$

Equation 2 shows that under illumination, the plasmonic particle adopts an electrochemical potential that is proportional to  $dT/dN$ . This quantity, which is determined only by the plasmon resonance spectrum and the heat flow from the particle to the substrate, provides the unique thermodynamic driving force in this system. It favors charge transfer to or from the particle that increases absorption, and thereby temperature, in order to lower the free energy. The factor  $dT/dN$  is largest on the steepest parts of the resonance spectrum; it is positive for irradiation on the blue side of the resonance, leading to a positive chemical potential for the electrons and hence a negative voltage. The reverse is



**Fig. 1. Absorption cross section spectrum of a Ag nanoparticle and the plasmoelectric effect.**

(A) Schematics: Spontaneous charge transfer to or from the nanoparticle is thermodynamically favored when the consequent spectral shift increases the absorption, raising the temperature. Irradiation on the blue side of the resonance leads to a negative charge on the particle; irradiation on the red side leads to a positive charge. (B) Calculated absorption cross section for a 20-nm-diameter Ag nanoparticle in vacuum with bulk carrier density  $n_{Ag}$  and carrier densities that are reduced or increased by 1%.

<sup>1</sup>Thomas J. Watson Laboratories of Applied Physics, California Institute of Technology, MC 128-95, Pasadena, CA 91125, USA. <sup>2</sup>Center for Nanophotonics, FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam, Netherlands.

\*Corresponding author. E-mail: haa@caltech.edu

observed for irradiation on the red side of the resonance. All these trends are exactly as observed in the experimental spectra of Fig. 2B: The surface potential vanishes at the peak of the resonance where  $dT/dN$  is zero; the largest (negative) potential is observed at a wavelength of 500 nm, coinciding with the steepest slope in the resonance spectrum of Fig. 2A.

Equation 1 can be used to quantitatively estimate the equilibrium plasmoelectric potential by writing  $F(N, T)$  as the sum of the free energies of electrons and phonons, using the well-known free energy functions of an electron and phonon gas (6, 7). The electronic term is composed of a contribution due to the chemical potential of the electrons that is directly given by the Fermi function, and an electrostatic contribution due to charging of the metal nanosphere; the phonon term is given by the Debye model. Taking the derivatives to  $N$  and  $T$ , we find analytical expressions for  $\mu(N, T)$  and  $S(N, T)$  that are then input into Eq. 2 (5).

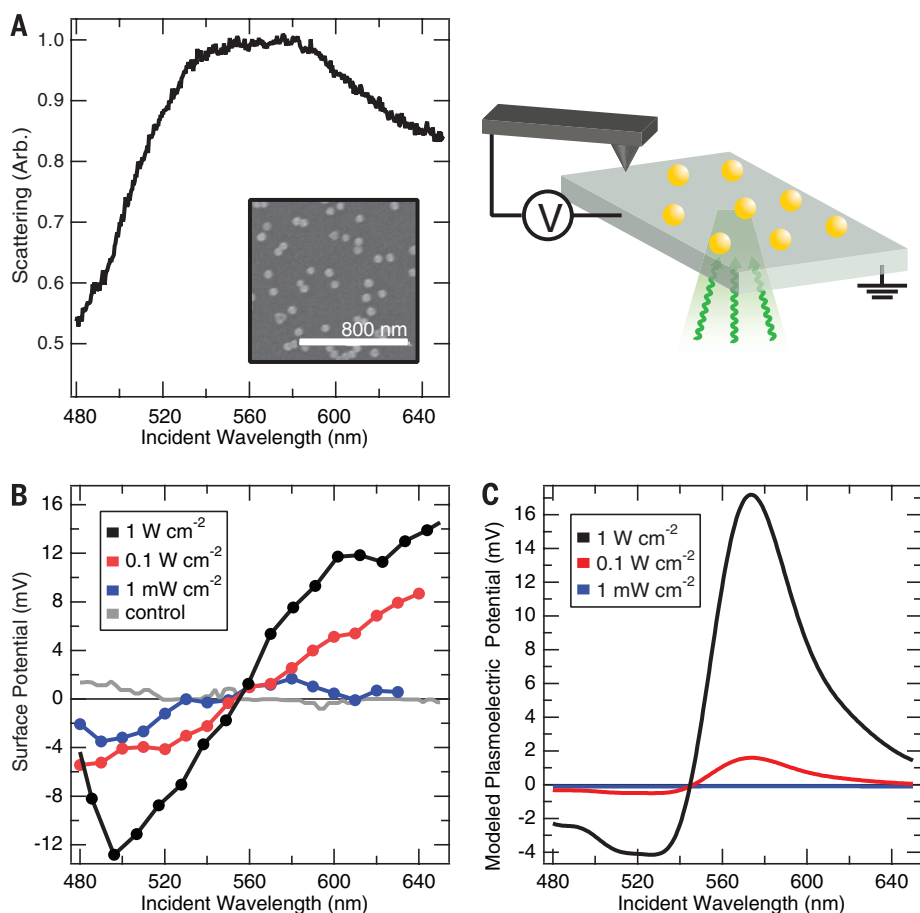
Before applying the model to the experimental geometry in Fig. 2, we first calculated the plasmoelectric potential (i.e., the electrochemical potential gained by the particle from the electron transfer induced by optical absorption) for a spherical 20-nm-diameter Ag nanoparticle in vacuum under monochromatic illumination. For these particles, analytical Mie theory can be used to calculate the absorption cross section spectrum,  $C_{\text{abs}}(\lambda, n)$ , for

a given electron density in the nanoparticle by taking into account the dependence of the bulk plasma frequency,  $\omega_p$ , in the complex dielectric function of the metal on carrier density,  $\omega_p \propto n^{1/2}$  (5). To calculate the nanoparticle temperature, we used a steady-state heat-flow model in which heat is dissipated from the nanoparticle by radiation. Figure 3A shows the calculated plasmoelectric potential of the Ag nanoparticle as a function of illumination wavelength at an incident flux of  $1 \text{ mW/cm}^2$ , under which the particle obtains a maximum temperature of  $\sim 400 \text{ K}$  (5). The model predicts a clear negative surface potential below the plasmon resonance and a positive one, up to 150 mV, above it. The asymmetry in the plasmoelectric potential below and above the resonance wavelength is due to the intrinsic nonresonant interband absorption in the metal. Figure 3B shows the corresponding relative absorption increase for the Ag nanoparticle, which ranges up to  $2.5 \times 10^{-5}$ .

The model for the simple geometry in Fig. 3 describes the key factors in the plasmoelectric effect: an increase in carrier density under illumination at wavelengths shorter than the resonance peak, inducing a negative plasmoelectric potential and enhanced absorption relative to the neutral particle. Similarly, radiation at longer wavelengths induces a lower carrier density, a positive plasmoelectric potential, and enhanced absorption. These results demonstrate that an excited

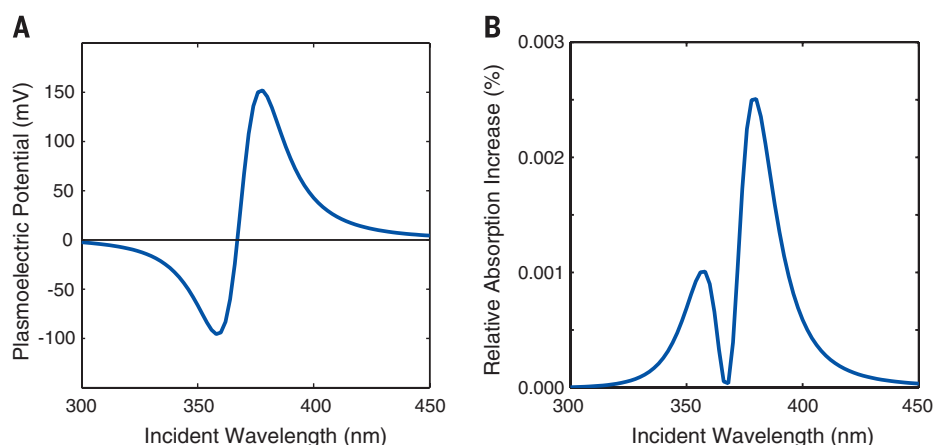
plasmonic resonator behaves as a heat engine that can convert absorbed off-resonant optical power into a static electrochemical potential.

Next, we used the model to calculate the wavelength-dependent and power-dependent plasmoelectric potential for the experimental geometry in Fig. 2, a spherical 60-nm-diameter Au particle on an ITO/glass substrate. We calculated the factor  $dT/dN$  in Eq. 2 using finite-difference time domain (FDTD) simulations of the absorption spectra for an Au particle on an ITO/glass substrate to take into account radiative damping from the substrate not captured by simple Mie theory and a one-dimensional model for heat conduction into the substrate (5). Because of variations in interparticle coupling and clustering [see scanning electron microscopy (SEM) image in Fig. 2A] and possible adsorption of water onto the particle-ITO interface, the ensemble dielectric environment is complex. The broadened, redshifted scattering spectrum in Fig. 2A is evidence for some particle aggregation. In the FDTD simulations, we modeled these effects by assuming a background index of  $n = 1.4$  for the medium above the ITO substrate, such that the absorption spectrum matched the experimentally observed spectrum. As can be seen in Fig. 2C for the high-power data, the modeled trends correspond well with the experimental trends: The modeled minimum potential occurs at 530 nm (experimental: 500 nm), the modeled zero potential

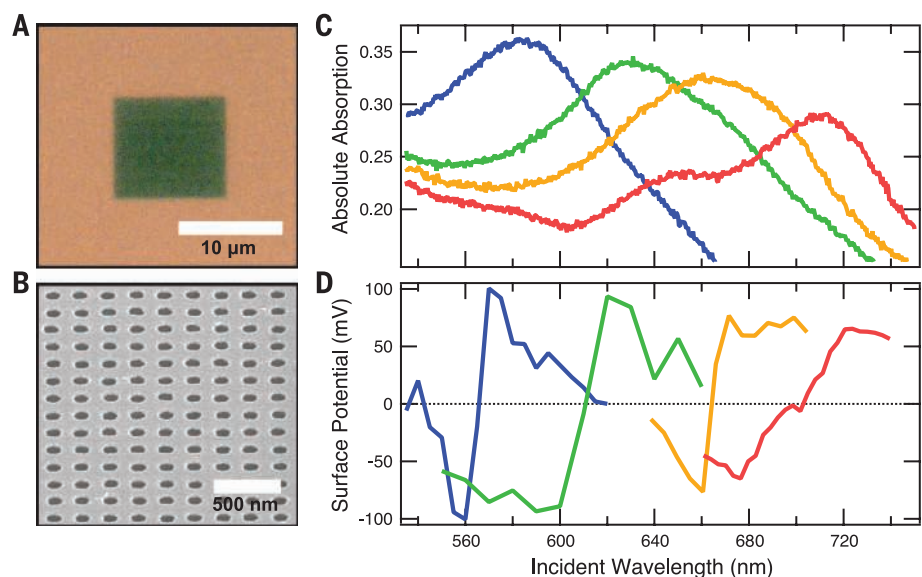


**Fig. 2. Plasmoelectric effect on dense Au nanoparticle arrays on ITO/glass.** (A) Dark-field scattering spectrum of 60-nm-diameter Au nanoparticles on ITO/glass. The inset shows a SEM image of the nanoparticle array. (B) KPFM measurements of the surface potential as a function of illumination wavelength (15-nm bandwidth) for three different illumination intensities. The surface potential of a flat region of ITO/glass adjacent to the nanoparticle array was monitored during scanned monochromatic illumination (see schematic geometry). A control measurement of an ITO/glass substrate without nanoparticles ( $1 \text{ W/cm}^2$ ) is also plotted (gray). (C) Modeled plasmoelectric potential for 60-nm-diameter Au nanoparticles on ITO/glass for the three illumination intensities in (B).





**Fig. 3. Modeled plasmoelectric response for Ag nanoparticles.** A 20-nm-diameter Ag particle in vacuum is illuminated with monochromatic light ( $1 \text{ mW/cm}^2$ ). (A) Plasmoelectric potential and (B) relative absorption increase as a function of incident wavelength.



**Fig. 4. Plasmoelectric effect on Au nanohole arrays on glass.** (A) Bright-field optical microscope image showing a 200-nm-pitch array of 100-nm-diameter holes in a 20-nm-thick Au film on glass. The dark green color directly shows the strong absorption in the hole array. (B) SEM image of a 200-nm-pitch hole array imaged under  $52^\circ$  with respect to the normal. (C) Measured optical absorption spectra for hole arrays with a pitch of 175 nm (blue), 225 nm (green), 250 nm (yellow), and 300 nm (red) ( $\text{NA} = 0.8$ ). (D) Surface potential as a function of excitation wavelength measured using KPFM at  $100 \text{ mW/cm}^2$  for the four-hole arrays in (C) ( $\text{NA} = 0.25$ ).

occurs at 545 nm (experimental: 560 nm), and a large positive potential is observed for wavelengths above the resonance, both in model and experiment. The extent of the measured potential to wavelengths up to 640 nm in Fig. 2B is in agreement with the long-wavelength tail in the spectrum of Fig. 2A. At 555 nm, the modeled particle temperature peaks at 308 K, 15 K above ambient (5), in good agreement with other experimental observations (8).

To aid interpretation of our findings, we comment briefly on other mechanisms for generat-

ing photopotentials with metals. A thermoelectric effect is several orders of magnitude weaker ( $\sim \mu\text{V K}^{-1}$ ) than the observed potentials (9). Hot carrier-induced effects would require rectifying contacts, which are not present in our geometry. Moreover, both of these effects would not result in a bisignated signal (10). "Plasmon drag" or similar direct photon-to-electron momentum transfer mechanisms on Au colloids would not produce a bisignated signal, nor would the hole arrays described below under normal-incidence irradiation (11). Future work in this

area will benefit from further insight into the microscopic mechanisms that contribute to the observed effect.

To demonstrate the generality of the observed plasmoelectric effect, we measured the plasmoelectric potential on arrays of subwavelength holes in an optically thin Au film on a glass substrate. In these microarrays, light is absorbed by the resonant excitation of interacting surface plasmon polaritons that propagate between the holes on the Au-air and Au-glass interface (5, 12). The absorption spectrum of a single array is determined by the coherent interaction of plasmons within the array, and the study of such arrays thus enables probing the plasmoelectric effect on a single absorber geometry. Moreover, the metal hole array architectures provide a first step toward future plasmoelectric integrated circuitry in which plasmoelectric power is harvested.

Figure 4, A and B, shows a fabricated  $10 \mu\text{m}$  by  $10 \mu\text{m}$  hole array composed of 100-nm-diameter holes in a square array with a pitch of 200 nm, made in a 20-nm-thick Au film on a glass substrate. Figure 4C shows the measured absorption spectra of hole arrays with a pitch of 175 to 300 nm (5). Increasing the pitch leads to a clear redshift in the absorption spectrum. Figure 4D shows the corresponding measured surface potential as a function of the wavelength of illumination ( $100 \text{ mW cm}^{-2}$ ), as measured with KPFM. A transition from negative to positive surface potential is seen for each array geometry. Also, the measured potential spectra show a clear redshift with increasing array pitch, in accordance with the trend in the absorption spectra. The maximum plasmoelectric potential observed for each measurement shows a gradually decreasing trend with increasing pitch, also in agreement with the trend in the absorption spectra. A slight difference is observed between the zero-crossing in the potential measurements and the peak in the corresponding absorption spectrum, which is attributed to the difference in numerical aperture (NA) of the microscope objective used for the two measurements (5). All surface potential spectra show clear minima and maxima and then decrease as the wavelength is tuned further away from the resonance, as expected. The measurements in Fig. 4D show a plasmoelectric potential as large as  $\pm 100 \text{ mV}$  and demonstrate the generality of the plasmoelectric effect for arbitrary plasmonic absorbers.

The observed plasmoelectric phenomenon takes advantage of the remarkable spectral tailorability of plasmonic nanostructures and can be extended to a variety of material systems, absorber geometries, and radiation environments. Plasmoelectric devices may enable the development of entirely new types of all-metal optoelectronic devices that can convert light into electrical energy by replacing the usual function of doped semiconductors with metal nanostructures that are optically excited off-resonance.

#### REFERENCES AND NOTES

1. J. A. Schuller et al., *Nat. Mater.* **9**, 193–204 (2010).
2. L. Novotny, N. van Hulst, *Nat. Photonics* **5**, 83–90 (2011).

3. C. Novo, A. M. Funston, A. K. Gooding, P. Mulvaney, *J. Am. Chem. Soc.* **131**, 14664–14666 (2009).
4. M. Nonnenmacher, M. P. O'Boyle, H. K. Wickramasinghe, *Appl. Phys. Lett.* **58**, 2921 (1991).
5. See supplementary materials on Science Online.
6. N. W. Ashcroft, N. D. Mermin, *Solid State Physics* (Harcourt, New York, 1976).
7. B. M. Askerov, S. Figarova, *Thermodynamics, Gibbs Method and Statistical Physics of Electron Gases* (Springer, New York, 2010).
8. M. G. Cerruti et al., *Anal. Chem.* **78**, 3282–3288 (2006).
9. F. J. Blatt, P. A. Schroeder, C. F. Foiles, *Thermoelectric Power of Metals* (Plenum, New York, 1976).
10. M. W. Knight, H. Sobhani, P. Nordlander, N. J. Halas, *Science* **332**, 702–704 (2011).
11. N. Noginova, A. V. Yakim, J. Soimo, L. Gu, M. A. Noginov, *Phys. Rev. B* **84**, 035447 (2011).
12. W. L. Barnes, W. A. Murray, J. Dintinger, E. Devaux, T. W. Ebbesen, *Phys. Rev. Lett.* **92**, 107401 (2004).

#### ACKNOWLEDGMENTS

Supported by U.S. Department of Energy (DOE) Office of Science grant DE-FG02-07ER46405 (M.S. and H.A.A.), an NSF Graduate Research Fellowship (A.B.), and experimental facilities of the DOE "Light-Material Interactions in Energy Conversion" Energy Frontier Research Center grant DE-SC0001293. Work at AMOLF is part of the research program of the Foundation for Fundamental Research on Matter [which is financially supported by the Netherlands Organization for Scientific Research (NWO)] and by the European

Research Council. The data are archived in the laboratory of H.A.A. and A.P. We thank E. Kosten, V. Brar, D. Callahan, M. Deceglie, A. Leenheer, J. Fakonas, R. van Roij, B. M. Mulder, and H. J. Bakker for helpful discussions.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6211/828/suppl/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S9  
References (13–36)

7 July 2014; accepted 16 October 2014  
Published online 30 October 2014;  
10.1126/science.1258405

## INTERFACIAL WATER

# The structure of interfacial water on gold electrodes studied by x-ray absorption spectroscopy

Juan-Jesus Velasco-Velez,<sup>1</sup> Cheng Hao Wu,<sup>1,2</sup> Tod A. Pascal,<sup>3</sup> Liwen F. Wan,<sup>3</sup> Jinghua Guo,<sup>4,5</sup> David Prendergast,<sup>3</sup> Miquel Salmeron<sup>1,6\*</sup>

The molecular structure of the electrical double layer determines the chemistry in all electrochemical processes. Using x-ray absorption spectroscopy (XAS), we probed the structure of water near gold electrodes and its bias dependence. Electron yield XAS detected at the gold electrode revealed that the interfacial water molecules have a different structure from those in the bulk. First principles calculations revealed that ~50% of the molecules lie flat on the surface with saturated hydrogen bonds and another substantial fraction with broken hydrogen bonds that do not contribute to the XAS spectrum because their core-excited states are delocalized by coupling with the gold substrate. At negative bias, the population of flat-lying molecules with broken hydrogen bonds increases, producing a spectrum similar to that of bulk water.

**M**any important processes in electrochemical reactions, such as ion desolvation and charge transfer, occur in the so-called electrical double layer (EDL), also known as the Helmholtz layer (*1*). This layer, containing solvated ions and solvent molecules, is typically one to tens of nanometers thick, depending on solute concentration. An important characteristic of the EDL is the presence of a strong electric field perpendicular to the electrode surface that is believed to play a crucial role in determining its structure.

Among liquids, aqueous solutions play a vital role in chemistry, biology, and materials sciences. The strong dipole of the water molecule and its hydrogen-bonding network gives rise to distinctive properties, including the high boiling point

and high specific heat capacity. It is expected therefore that perturbations of the local hydrogen bonding network near surfaces (*2, 3*), together with strong electric fields, will influence the structure of the first molecular layers. Understanding the structure and dynamics of these interface layers is of fundamental importance and the object of our study.

A host of techniques have been used to study the EDL, including grazing incidence x-ray scattering (*4*) that provides crystallographic information and electronic density profiles, and vibration spectroscopies such as infrared (IR) and Raman (*5, 6*), together with nonlinear second-order sum frequency generation (SFG), which have been used to study the vibrational nature of water molecules at interfaces (*7*). Using SFG at the glass-water interface at different pH values, it was concluded that protonation and deprotonation of oxide sites led to surface charges and electric fields that align the interfacial water molecules (*8–10*).

In contrast to these techniques, x-ray absorption spectroscopy (XAS) is element-specific and provides information about the electronic structure around the excited atom, which is sensitive to the local structure and chemical environment (*11*). XAS has been extensively used to study the

structure of water by analyzing the oxygen K-edge spectra (*12, 13*). XAS can be performed by collecting either photons or secondary electrons generated by the decay of the core holes created by the absorbed x-rays. The first detection mode, called total fluorescence yield (TFY) (*14–17*), is a bulk-sensitive technique, because the mean free path of soft x-rays is in the micrometer range. When collecting secondary electrons (the so-called total electron yield mode, TEY), the information is surface sensitive because of the short mean free path of electrons in condensed matter; on the order of nanometers for energies in the 100-eV range.

Our system consists of an electrochemical flow cell with an ultrathin Si<sub>3</sub>N<sub>4</sub> membrane window (100 nm thick) that separates the liquid medium inside the cell from the high-vacuum environment of Beamline 8.0.1 in the Advanced Light Source at the Lawrence Berkeley National Laboratory (Fig. 1). The cell contains reference and counter electrodes for electrochemical experiments (*18, 19*). The electrode under study is a ~20-nm Au film evaporated on the internal side of the Si<sub>3</sub>N<sub>4</sub> membrane, in contact with a flowing aqueous 10 μM solution of NaCl. The film is continuous and fully covers the membrane surface, as shown by atomic force microscopy (AFM) imaging (fig. S1) (*20*). Gold was chosen because its inert nature makes it possible to explore a wide range of electrostatic potentials without inducing electrolysis or other electrochemical reactions that could alter the structure and composition of its surface. The lack of reactivity or ion adsorption was confirmed by the absence of peaks in the cyclic voltammetry curves (fig. S2) (*20*). Because the effective attenuation length of electrons with energy smaller than 530 eV (from the O K-edge excitation of water) is estimated to be between 5 and 10 Å in liquid water (*21, 22*), the TEY signal should originate from the first two or three water layers. This was confirmed by theoretical calculations as discussed below.

Figure 2A shows the O K-edge XAS spectra of water at open circuit potential in TFY (curve a), and TEY detection modes (curves b and c). Curve a corresponds to the well-known oxygen K-edge spectrum of bulk water, which is divided into three regions: pre-edge region, around 535 eV (I); main edge region, around 537 eV (II); and post-edge region, around 540 eV (III). The pre-edge peak at 535 eV is characteristic of the liquid phase

<sup>1</sup>Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. <sup>2</sup>Department of Chemistry, University of California, Berkeley, CA 94720, USA. <sup>3</sup>Joint Center for Energy Storage Research, The Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. <sup>4</sup>The Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. <sup>5</sup>Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064, USA. <sup>6</sup>Department of Materials Science and Engineering, University of California, Berkeley, CA 94720, USA.

\*Corresponding author. E-mail: mbsalmeron@lbl.gov



and originates from unsaturated donor (or dangling) hydrogen bonds (11, 12, 23). The main edge is similarly related to the population of molecules with unsaturated hydrogen bonds, while the post-edge region is associated with the fully saturated hydrogen bonding network. In bulk ice, the pre- and main-edge peaks are weak because all of the hydrogen atoms participate in hydrogen bonds. In liquid water, the pre-edge peak is relatively prominent, while the post-edge is weaker compared to ice (12, 24). As can be seen in the TEY spectrum (curve b), no spectral feature is visible at 535 eV, which might be interpreted as indicative of a much lower concentration of broken hydrogen bonds at the gold-water interface compared to bulk water. In fact, as we shall see below, the contrary is true.

To gain insights into the measurements, we performed *ab initio* molecular dynamics (AIMD) simulations, to explore the interfacial structure of water near gold. We initiated our AIMD simulation from a pre-equilibrated box comprising 150 water molecules interfaced with six layers of gold atoms oriented to expose the (111) surface (Fig. 2B). To interpret the XAS measurements, we used a smaller model of the same interface and sampled five evenly spaced snapshots from the last 5 ps of AIMD trajectory. The calculation of the XAS was done with the constrained-occupancy density functional theory within the excited electron and core-hole (XCH) approach (23), which has been shown to be efficient in predicting the XAS of complex molecular systems (25, 26).

Analysis of the AIMD trajectory revealed that 49% of the interfacial water molecules (i.e., water molecules within ~1 nm of the electrode) lie flat on the surface with both hydrogen bonds saturated, denoted as double-donor (DD) species (Fig. 3). These molecules are characterized by an XAS spectrum with no substantial pre-edge feature. Another 49% of the interfacial water molecules have one broken hydrogen bond, either parallel to the surface ( $SD^{\parallel}$ , 18%) or perpendicular ( $SD^{\perp}$ ) with one hydrogen atom pointing toward the gold surface (31%). This population of broken H-bond molecules is substantially higher than the 22% found in bulk water.

We find that although the XAS of the  $SD^{\parallel}$  molecules resembles that of SD molecules in the bulk, the XAS of the  $SD^{\perp}$  molecules does not present a pre-edge feature. Similarly, water molecules with two dangling hydrogen bonds (ND species) are oriented with both hydrogen atoms pointing toward the gold surface and thus do not contribute to the pre-edge feature.

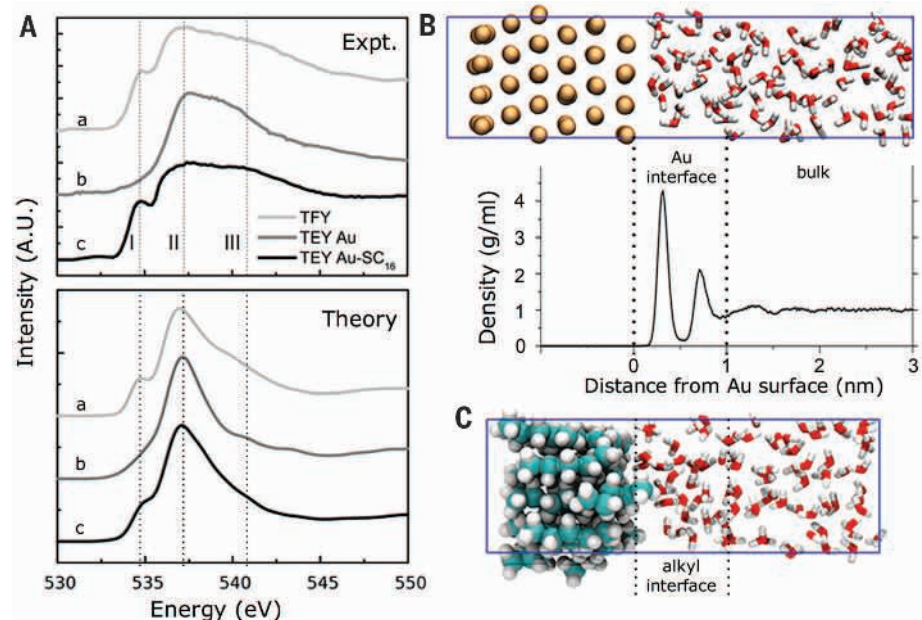
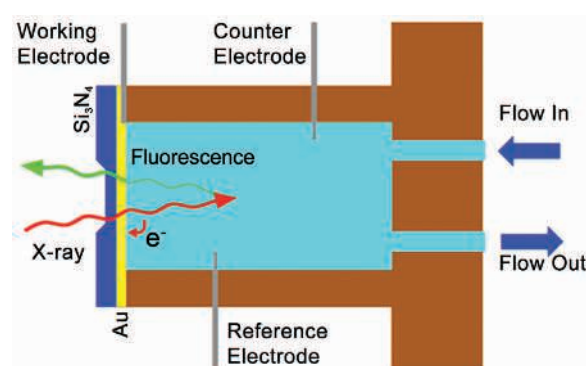
The suppression of the pre-edge of perpendicular water molecules with broken hydrogen bonds at the gold surface is a purely electronic effect. The core-excited state resulting from the absorption of a 535-eV x-ray photon is coupled to the gold band structure and delocalizes over the surrounding gold atoms (fig. S3) (20). This delocalization reduces the overlap of the core-excited electron with the 1s core hole at the excited oxygen atom, thereby reducing the transition probability and thus the peak intensity. Further, because

the pre-edge peak arises from antibonding  $\sigma^*$  states centered on the hydrogen atom and extending directionally along the OH bond vector (27), coupling to the gold depends strongly on the orientation of the water molecule and affects primarily the water molecules in the first layer. The water molecules with broken hydrogen bonds in the second layer also have a reduced pre-edge intensity, although the reduction is only about 30% that of the first layer. The net result is that the XAS obtained from sampling only the interfacial water molecules show a substantially reduced

pre-edge feature. The good agreement with the TEY spectrum confirmed the high interfacial sensitivity of these measurements.

Our AIMD results, which show a large number of broken hydrogen bonds in the interfacial region, are consistent with previous simulations reporting substantial populations of interfacial water molecules oriented with hydrogen atoms pointing toward the Au surface (28, 29). It is at odds, however, with conclusions based on studies of interfacial water at cryogenic temperatures, which may not represent the room-temperature liquid-solid

**Fig. 1. Schematic of the electrochemical cell.** A  $Si_3N_4$  membrane (~100 nm thick) separates the liquid from the vacuum region of the synchrotron x-ray source. The gold electrode is a thin (20 nm) film evaporated on the back side of the membrane. Detection of x-ray absorption was done by collecting the fluorescence emission on the vacuum side and also by collecting secondary electrons at the gold electrode. The first method yields bulk sensitive spectra. The second is sensitive to the liquid in the double layer because of the short mean free path of electrons (~1 nm).



**Fig. 2. Water/gold interface at open circuit potential.** (A) Experimental and theoretical XAS at the O K-edge of water near the gold electrode, measured under open circuit in the liquid flow cell (Fig. 1). Curves marked (a) correspond to the total fluorescence yield (TFY) spectrum and calculated XAS spectrum for bulk water. Curves marked (b) correspond to the total electron yield (TEY) spectrum and calculated XAS spectrum for the interfacial water molecules next to the gold electrode. Curves marked (c) correspond to the TEY spectrum of water after covering the gold surface with a hexadecanethiol ( $C_{16}SH$ ) monolayer and calculated XAS spectrum for interfacial water molecules next to an alkyl surface, respectively. Snapshots from AIMD simulations of (B) Au-water, with associated mass density as a function of distance from the surface, and (C)  $C_{16}SH$ -water interface, with the gold atoms omitted in the latter case for computational efficiency. The water column separating the Au electrodes is large enough that the water density fluctuations are smoothed beyond the first few layers. Only half of the water column space is shown.

interface. For example, STM and infrared spectroscopy studies of a nonwetting bilayer film have revealed fully saturated H-bonds at the Au(111) surface (30). By contrast, calculations (31, 32) and scanning tunneling microscopy (STM) studies of water on more reactive surfaces, such as Ru and Pd, indicate the formation of hexagonal

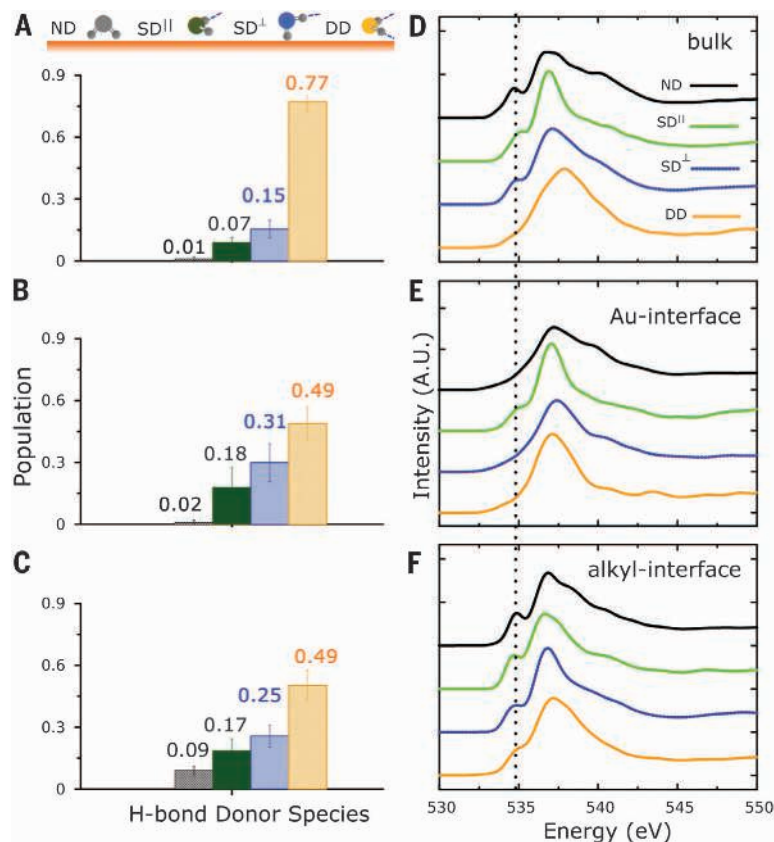
structures in the first layer of water, where the molecules lie flat on the surface and bind to the metal through the oxygen lone-pair electrons in geometries that maximize the number of hydrogen bonds to other water molecules (33).

To verify that the suppression of the pre-edge in the XAS of interfacial water molecules is caused

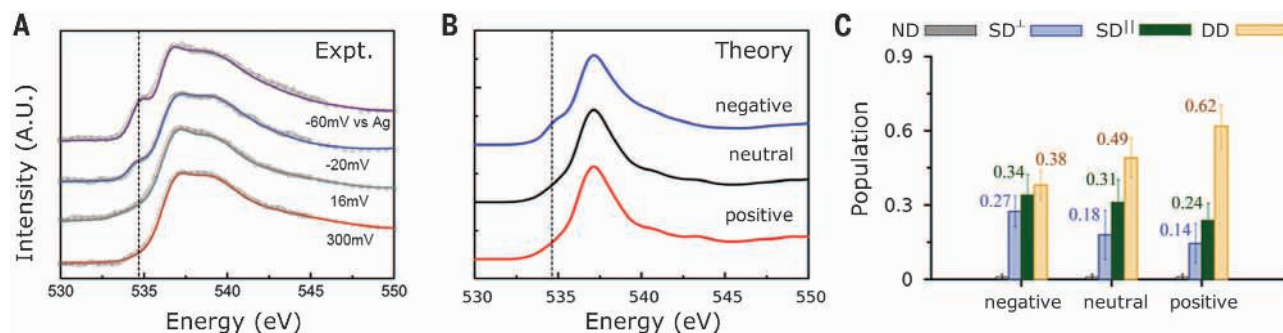
by coupling to the gold electronic structure, we performed an experiment with the gold surface covered with a monolayer of hexadecanethiol ( $C_{16}H_{33}SH$ ) that rendered the surface hydrophobic. AFM images of this surface revealed the absence of defects exposing the gold substrate within the  $\sim 1$ -nm resolution of the technique (34). Our AIMD calculations revealed an increase in the SD and ND populations at the interface (42 and 9% respectively, relative to the bulk (22 and 1%, respectively; Fig. 3, A and C). The large population of SD and ND species from our AIMD simulations is consistent with vibrational SFG results by Shen *et al.*, which showed a substantial amount of dangling hydrogen bonds in water at the alkyl-water interface (7). The TEY XAS result (curve c in Fig. 2) showed a prominent peak at 535 eV characteristic of liquid water, whereas the TFY XAS remained the same as in the case of bare Au/water. Our calculations indicate that the low-energy core-excited states of the interfacial SD and ND molecules are no longer coupled with the (now distant) gold and are highly localized on the excited molecules, therefore producing the pre-edge absorption feature seen in the bulk.

Because water molecules have a strong dipole moment, it is expected that they will respond to external electrical fields (35). Infrared spectroscopy results have indeed shown a bias-dependent orientation of the molecules by monitoring the intensity of O-H and H-O-H vibrational modes as a function of applied bias (5). We thus investigated the influence of electric fields on the interfacial water structure in the EDL. To separate the large Faradaic current reaching the sample from the electrolyte (typically microamperere or hundreds of nanoamperere) and the much smaller TEY XAS current (500 pA to 10 nA), we modulated the incoming x-ray beam with a piezo-actuated chopper (shown schematically in fig. S5) (20) and measured the modulated TEY current with a lock-in amplifier. In this manner, XAS spectra were obtained at bias voltages ranging from  $-60$  to  $300$  mV relative to a Ag pseudo-reference electrode (20).

The results are shown in Fig. 4. At positive bias, the pre-edge peak in the O K-edge spectrum



**Fig. 3. Populations of water molecules of various orientations obtained from AIMD simulations and calculated XAS.** In AIMD simulation, the water molecules were sampled: (A) from bulk water, (B) at the Au interface, and (C) at the alkyl interface. Schematics of the nondonor (ND, black), single-donor parallel and perpendicular to the surface ( $SD^{\parallel}$  green,  $SD^{\perp}$  blue), and double-donor (DD, yellow) molecules shown at the top. Same color code is used for the calculated XAS from water molecules sampled from bulk water (D), at the Au interface (E), and at the alkyl interface (F).



**Fig. 4. The potential-dependence of water XAS spectra and calculated hydrogen-bonding populations.** (A) TEY O K-edge XAS of water collected at the Au electrode under different potentials versus a Ag quasi-reference electrode. The point of zero charge (PZC) is found to be at 80 mV (fig. S6). At positive bias, the pre-edge peak at 535 eV characteristic of dangling H-bonds is negligible, whereas it is prominent under negative bias. (B) Calculated O K-edge XAS. The pre-edge peak at negative bias is the result of an increase in the population of single-donor ( $SD^{\parallel}$ ) molecules lying parallel to the surface with core-excited states that are decoupled from the gold. (C) Population of H-bonded water molecules as a function of the Au surface charge state.



was negligible, as in the case without applied bias, but under negative bias, the pre-edge feature became prominent. Because no electrochemical reactions or ion adsorption take place at these bias voltages (fig. S2) (20), a simple interpretation of the observations is that under negative bias, the electric field favors an orientation of the water molecules with their H atoms toward the gold surface, which increases the number of dangling hydrogen bonds. This hypothesis was confirmed by our AIMD simulations, which show an enhanced DD population at positive bias and an enhanced SD population at negative bias (Fig. 4). Indeed, at negative bias, the change in the orientation of the water molecules greatly disrupts the hydrogen bonding network in the interfacial layer (fig. S4) (20). This gives rise to an increased SD<sup>II</sup> population that causes a prominent pre-edge feature in the XAS.

Because XAS is element-specific, our experimental method opens the way for further studies of the structure and chemistry of solvent and solute species in the interfacial layers close to an electrode (~1 nm), and in the presence of electric fields. The complexity of the interfacial molecular rearrangements underscores the need for accurate and efficient first-principles calculations to aid interpretation and, as in the present case, to uncover previously unknown physics related to the strong coupling of x-ray excited states to surface electronic structure. This combined experimental and theoretical approach is essential for a fundamental understanding of electrochemical reactions, with applications to electrocatalysis, photochemistry, and energy storage, among others.

## REFERENCES AND NOTES

1. A. J. Bard, L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications* (Wiley, New York, 1980).
2. G. Cicero, J. C. Grossman, E. Schwegler, F. Gygi, G. Galli, *J. Am. Chem. Soc.* **130**, 1871–1878 (2008).
3. D. Chandler, *Nature* **437**, 640–647 (2005).
4. P. Fenter et al., *J. Colloid Interface Sci.* **225**, 154–165 (2000).
5. K. Ataka, T. Yotsuyanagi, M. Osawa, *J. Phys. Chem.* **100**, 10664–10672 (1996).
6. M. Fleischmann, P. J. Hendra, I. R. Hill, M. E. Pemble, *J. Electroanal. Chem.* **117**, 243–255 (1981).
7. Y. R. Shen, V. Ostroverkhov, *Chem. Rev.* **106**, 1140–1154 (2006).
8. L. Zhang, C. Tian, G. A. Waychunas, Y. R. Shen, *J. Am. Chem. Soc.* **130**, 7686–7694 (2008).
9. V. Ostroverkhov, G. A. Waychunas, Y. R. Shen, *Phys. Rev. Lett.* **94**, 046102 (2005).
10. J. Sung, L. Zhang, C. Tian, Y. R. Shen, G. A. Waychunas, *J. Phys. Chem. C* **115**, 13887–13893 (2011).
11. A. Nilsson et al., *J. Electron Spectrosc. Relat. Phenom.* **177**, 99–129 (2010).
12. H. Bluhm, D. F. Ogletree, C. S. Fadley, Z. Hussain, M. Salmeron, *J. Phys. Condens. Matter* **14**, L227–L233 (2002).
13. A. Nilsson, L. G. M. Pettersson, *Surf. Sci. Rep.* **55**, 49–167 (2004).
14. J.-H. Guo et al., *Phys. Rev. Lett.* **89**, 137402 (2002).
15. O. Fuchs et al., *Nucl. Instruments Methods A* **585**, 172–177 (2008).
16. T. Tokushima et al., *J. Chem. Phys.* **136**, 044517 (2012).
17. J.-H. Guo et al., *Phys. Rev. Lett.* **91**, 157401 (2003).
18. P. Jiang et al., *Electrochem. Commun.* **12**, 820–822 (2010).
19. A. Braun et al., *J. Phys. Chem. C* **116**, 16870–16875 (2012).
20. See supplementary materials in Science Online.
21. S. Ghosal et al., *Science* **307**, 563–566 (2005).
22. M. A. Brown, M. Faubel, B. Winter, *Annu. Rep. Prog. Chem. Sect. C* **105**, 174 (2009).
23. D. Prendergast, G. Galli, *Phys. Rev. Lett.* **96**, 215502 (2006).
24. U. Bergmann et al., *Phys. Rev. B* **66**, 092107 (2002).
25. A. H. England et al., *Chem. Phys. Lett.* **514**, 187–195 (2011).
26. W. S. Drisdell et al., *J. Am. Chem. Soc.* **135**, 18183–18190 (2013).
27. P. Wernet et al., *Science* **304**, 995–999 (2004).
28. G. Cicero, A. Calzolari, S. Corni, A. Catellani, *J. Phys. Chem. Lett.* **2**, 2582–2586 (2011).
29. R. Nadler, J. F. Sanz, *J. Chem. Phys.* **137**, 114709 (2012).
30. D. Stacchiola et al., *J. Phys. Chem. C* **113**, 15102–15105 (2009).
31. P. J. Feibelman, *Science* **295**, 99–102 (2002).
32. A. Michaelides, A. Alavi, D. A. King, *J. Am. Chem. Soc.* **125**, 2746–2755 (2003).
33. M. Tatarkhanov et al., *J. Am. Chem. Soc.* **131**, 18425–18434 (2009).
34. G. Liu, M. Salmeron, *Langmuir* **10**, 367–370 (1994).
35. M. F. Toney et al., *Nature* **368**, 444–446 (1994).

## ACKNOWLEDGMENTS

This work was supported by the Office of Basic Energy Sciences (BES), Division of Materials Sciences and Engineering, of the U.S. Department of Energy (DOE) under contract no. DE-AC02-05CH11231 (through the Chemical and Mechanical Properties of Surfaces, Interfaces and Nanostructures program). J.-J.V.-V. acknowledges financial support from the Alexander von Humboldt Foundation, Germany. C.H.W. acknowledges the Advanced Light Source (ALS) Doctoral Fellowship in Residence. Theory and

simulations by T.A.P., L.F.W., and D.P. were supported by the Joint Center for Energy Storage Research, an Energy Innovation Hub funded by the U.S. DOE and facilitated by a user project at the Molecular Foundry. Computations were performed with the computing resources of the National Energy Research Scientific Computing Center (NERSC). The ALS and Molecular Foundry (supported by DOE-BES) and NERSC (supported by DOE-Advanced Scientific Computing Research) are DOE Office of Science User Facilities, supported by the DOE Office of Science under contract no. DE-AC02-05CH11231. We thank C.-H. Chuang, B.-Y. Wang, D. Zhang, X. Feng, and M. W. West for support at the beamline. We also thank J. Zhang for help with AFM imaging and C. Das Pemmaraju, C. Schwartz, P. Ross, J. Colchero, G. Thornton, H. Fang, and S. Harris for useful discussions.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6211/831/suppl/DC1  
Methods

Figs. S1 to S6

References (36–65)

31 July 2014; accepted 9 October 2014

Published online 23 October 2014;

10.1126/science.1259437

## STRONG BASES

# Directed ortho-meta'- and meta-meta'-dimetalations: A template base approach to deprotonation

Antonio J. Martínez-Martínez, Alan R. Kennedy, Robert E. Mulvey,\* Charles T. O'Hara\*

The regioselectivity of deprotonation reactions between arene substrates and basic metalating agents is usually governed by the electronic and/or coordinative characteristics of a directing group attached to the benzene ring. Generally, the reaction takes place in the ortho position, adjacent to the substituent. Here, we introduce a protocol by which the metalating agent, a disodium-monomagnesium alkyl-amide, forms a template that extends regioselectivity to more distant arene sites. Depending on the nature of the directing group, ortho-meta' or meta-meta' dimetalation is observed, in the latter case breaking the dogma of ortho metalation. This concept is elaborated through the characterization of both organometallic intermediates and electrophilically quenched products.

One of the most widely applied chemical reactions is metalation (1, 2), in which an inert carbon-hydrogen (C-H) bond is transformed to a more reactive carbon-metal (C-M) bond by a metalating agent. The fundamental importance and vast scope of metalation arise from the ubiquity of the C-H bond, which is one of the most abundant bonds found in nature and provides a rich sustainable entry point for the synthesis of aromatic chemicals, natural products, and organic-based materials. However, this abundance of C-H bonds poses a formidable challenge to synthetic chemists: How can metalation reactions be made regioselective—that is, to deprotonate specific C-H bonds without reacting with other C-H bonds in the same molecule? One answer arrived in “directed ortho-metalation” (DoM) (3–6), the seminal concept to

date in metalation chemistry. First introduced independently by Gilman (7) and Wittig (8), and then greatly extended by Beak (9), Mortier (10), Hoppe (11, 12), and Snieckus (13, 14), among others, DoM relies primarily on the substitution within the aromatic substrate undergoing the C-H to C-M transformation. To induce the DoM reaction, this substrate must contain a directing metalation group (DG) that can activate an adjacent (ortho) C-H bond toward metalation either by providing the incoming Lewis acidic metalating agent with a Lewis basic docking site (coordination effect) and/or weakening this bond through electron-withdrawing inductive properties (electronic effect) (Fig. 1). Depending on their relative coordinating and electron-withdrawing ability, DGs can be weak, moderate, or strong ortho-directors.

In general, the DoM concept applies irrespective of which metalating agent is used, whether it be an organolithium reagent or one of the new wave of bimetallic formulations [typified

WestCHEM, Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1XL, UK.

\*Corresponding author. E-mail: r.e.mulvey@strath.ac.uk (R.E.M.); charlie.ohara@strath.ac.uk (C.T.O.)

by Knochel's (TMP)MgCl·LiCl (15, 16); TMP is 2,2,6,6-tetramethylpiperidine]. Emphasizing this limited regioselectivity, Schlosser notes in his mas-

terwork on polar organometallic chemistry that “deviations from the ortho rule do exist but are scarce” (2). Beak and Snieckus in their seminal



Fig. 1. Mechanism of the substituent DoM.

Fig. 2. DG substituted arenes for screening with **2<sub>Bu</sub>**, highlighting their comparative ortho-directing ability.

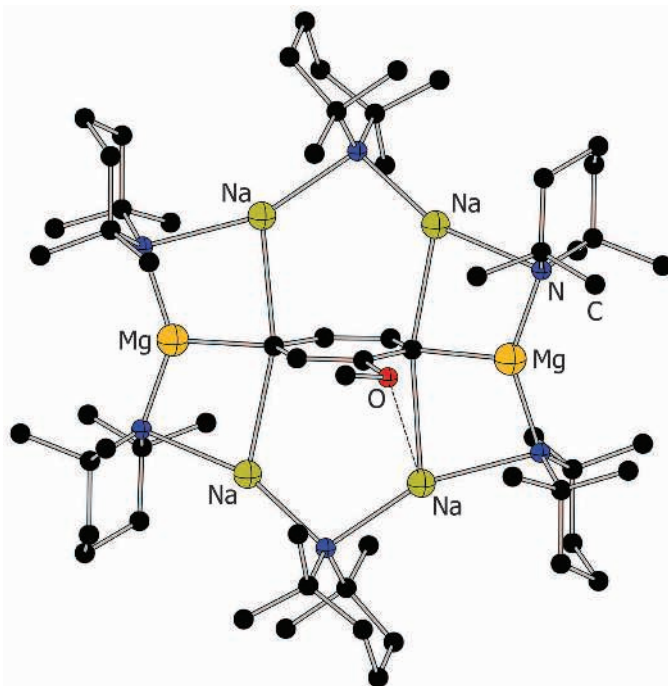
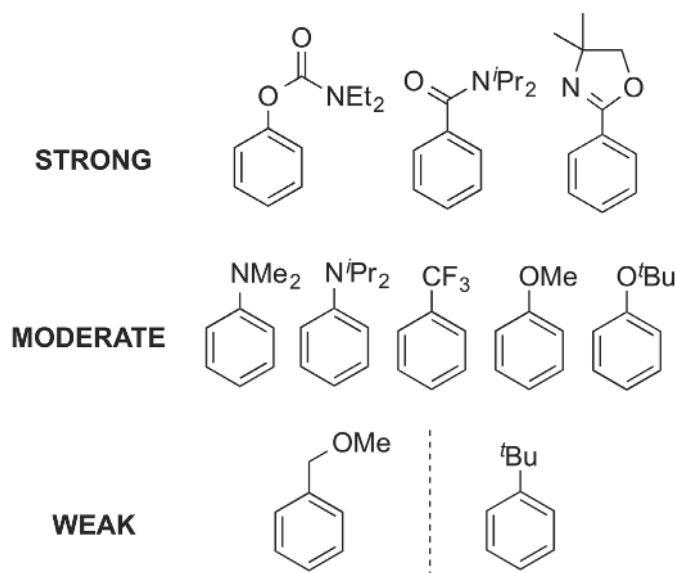


Fig. 3. Molecular structure of  $[\text{Na}_4\text{Mg}_2(\text{TMP})_6(\text{C}_6\text{H}_3\text{OMe-2,5})]$  **5a**. Hydrogen atoms and one disordered component of both one TMP ligand and the anisole dianion have been omitted for clarity. The dashed line illustrates a  $\text{Na}\cdots\text{O}$  contact.

perspective on the “complex-induced proximity effect” (CIPE) (6) explain how lithiation can occur remotely from the ortho position in special cases in which the target C-H bond is formally remote (through bond connections) but is conformationally in proximity (through space) to the lithium-DG substituent coordination point. Recently, this idea has been elegantly extended to cyclopalladation reactions of certain amines (17), hydrocinnamic acids, and toluenes (18) through use of conformationally locked substituents that effect meta-activation.

Here, we introduce a concept to extend the scope of regioselectivities attainable in aromatic C-H metalation chemistry. The innovation lies in controlling the regioselectivity through the structure of the metalating agent rather than the inherent DG characteristics of the aromatic ring substituents.

Recently, we established that the donor solvent-free structure of the potassium alkyl-amido magnesiate hexamer  $[\{\text{KMg}(\text{TMP})_2(^n\text{Bu})\}_6]$  **1<sub>Bu</sub>** is capable of metalating benzene, toluene, and naphthalene to produce the hexameric areneides  $[\{\text{KMg}(\text{TMP})_2(\text{areneide})\}_6]$  **1<sub>areneide</sub>** (19). This transformation provided strong evidence for a base template mechanism: The six butyl components of **1<sub>Bu</sub>** deprotonate regioselectively one C-H bond of six arene molecules to generate **1<sub>areneide</sub>** inverse crowns, in which the areneide anions are trapped within the retained hexameric ring framework of the base. Describing such compounds as inverse crowns acknowledges the topological but inverse relation to conventional crown ethers in which the Lewis basic (O heteroatom containing) host rings trap Lewis acidic (metal atom) guests. A donor solvent-free medium (methylcyclohexane) was essential for these templation reactions because without competition from lone pair coordinating solvents (such as ethers), the pre-inverse-crowns can grow uninterruptedly through polar but molecular (not infinite) aggregation.

With the sixfold symmetry of **1<sub>Bu</sub>** leading to similarly symmetrical inverse crown products, we thought it prudent to probe the metalating ability of sodium magnesiate “ $[\text{Na}_4\text{Mg}_2(\text{TMP})_6(^n\text{Bu})_2]$ ” **2<sub>Bu</sub>** in methylcyclohexane because its lower symmetry might promote different C-H deprotonating reactions. Promisingly, a preliminary reaction of **2<sub>Bu</sub>** with naphthalene produced an inverse crown comprising a 12-atom  $[\text{Na}(\text{TTHP})\text{Na}(\text{TMP})\text{Mg}(\text{TMP})_2]$  host ring **3** that trapped a 1,4-dideprotonated naphthalene ( $1,4\text{-C}_{10}\text{H}_6$ ), although this reaction was complicated by one TMP ligand losing methane to form TTHP (2,2,6-trimethyl-1,2,3,4-tetrahydropyridine) (19, 20). Although attempts to crystallize **2<sub>Bu</sub>** have been unsuccessful, we have characterized its hydrocarbon solution structure by means of several nuclear magnetic resonance (NMR) spectroscopic methods. Most compelling, diffusion-ordered spectroscopy (DOSY) and exchange spectroscopy (EXSY) NMR spectra show two *n*-butyl-containing species in solution, which can be attributed to the existence of a monomer-dimer equilibrium between “ $[\text{Na}_2\text{Mg}(\text{TMP})_3(^n\text{Bu})]$ ” and its dimer **2<sub>Bu</sub>** (figs. S4



to S6 and accompanying text). Equilibria of this type are common in polar organometallic chemistry (27). Also, inverse crowns of matching formula  $[\{Na_4Mg_2(TMP)_6\}(\text{areneide}^{2-})]$  are known to form from in situ mixtures of **2<sub>Bu</sub>** with benzene or toluene (22), whereas similar 12-atom ring motifs can be constructed by replacing magnesium with divalent metals of comparable size (Cr, Mn, or Fe) (23).

To cover the full range of weak to strong ortho-directors, we screened all the representative DoM substrates in Fig. 2 with **2<sub>Bu</sub>** prepared in methylcyclohexane solution by combining <sup>1</sup>BuNa (24), <sup>1</sup>Bu<sub>2</sub>Mg, and TMP(H) in a 2:1:3 ratio. We have developed a user-friendly, one-pot, glovebox-free strategy for the synthesis of **2<sub>Bu</sub>**, which is notably stable as a 0.1 M solution in methylcyclohexane for at least 1 month (25). We began the screening with anisole **4a**.

Anisole, an important scaffold in pharmaceuticals such as the analgesic Tramadol (26) and the antidepressant Moxifetin hydrogen maleate (27), has been a key benchmark reagent in enhancing understanding of DoM chemistry. Until now, it has only been possible to directly metalate anisole at one ortho-position (28–31). Reaction of anisole with **2<sub>Bu</sub>** in methylcyclohexane solution under reflux conditions for 12 hours resulted in 2,5-dimetalation; work-up with D<sub>2</sub>O produced 2,5-anisole-d<sub>2</sub> **6a** (Table 1, entry 1, 80%). To gain insight into the intermediate organometallic chemistry behind this transformation, we undertook an x-ray diffraction study of crystalline **5a** that revealed the new inverse crown  $[Na_4Mg_2(TMP)_6(C_6H_3OMe-2,5)]$  (Fig. 3). Although the x-ray data are of insufficient quality to discuss structural metrics, the metalation sites can be unequivocally assigned to an ortho- and opposite-facing meta'-position.

Mg atoms within the 12-atom metal-amide ring occupy these deprotonated sites consistent with dimagnesiumation of anisole. <sup>1</sup>H NMR spectroscopic data of both the isolated solid and reaction mixture show that the dimagnesium-tetrasodium hexanuclear ring is retained in *cyc*-C<sub>6</sub>D<sub>12</sub> solution (25). The observation of only one set of anisole resonances confirmed that double deprotonation has occurred regioselectively at the 2,5-positions of anisole, the sole regioisomer formed. By treating in situ solutions of **5a** with iodine/tetrahydrofuran (THF) or carbon dioxide then water, 1,4-diiodo-2-methoxybenzene (**6a'**) and 2-methoxyterephthalic acid (**6a''**) were isolated, respectively (entry 2).

Subjecting (methoxymethyl)benzene (**4b**), *tert*-butoxybenzene (**4c**), (trifluoromethyl)benzene (**4d**), *N,N*-diisopropylbenzamide (**4e**), 4,4-dimethyl-2-phenyl-4,5-dihydrooxazole (**4f**), and phenyl-*N*,*N*-diethyl-*O*-carbamate (**4g**) to **2<sub>Bu</sub>** also gave 2,5-dimetalated intermediates in excellent yields (25). Quenching them with electrophiles gave the respective tri-substituted arenes in good to excellent isolated yields (Table 1).

In contrast, reaction of **2<sub>Bu</sub>** with *N,N*-dimethylaniline (**7a**) and *N,N*-diisopropylaniline (**7b**) (Table 2) induced selective 3,5-metalation of the substrate evidenced through two singlets in a 2:1 integration ratio in the arene region of <sup>1</sup>H NMR spectra of organometallic intermediates **8a** and **8b**.

X-ray diffraction studies of crystalline of **8a** (91% yield) revealed a structure resembling **5a**, but with the aniline guest deprotonated in the 3,5-positions. This example of directed meta-meta' metalation (DmmM) completely overrides the conventional DoM effect observed in the anilines. **8a** and **8b** could be converted to their respective diiodo-compounds in good yields by treatment with iodine/THF (Table 2).

X-ray data (Fig. 4) reveal that the anilino-N atom in **8a** exhibits planar sp<sup>2</sup> hybridization, allowing conjugation of its lone pair into the aromatic system. Consequently, hindered rotation about the N-C<sub>ipso</sub> bond effectively results in the methyl groups offering permanent steric protection to the ortho-positions, thus altering the metalation regioselectivity from 2,5- to 3,5- for steric reasons. To

pursue this idea further, we treated **1** with the non-DoM-directing *t*-butylbenzene (**7c**). Unlike the anilines, ortho-protection is not fixed, but we posited that the additional methyl group adjacent to the ipso-C may provide a similar ortho-protection. Quenching the reaction with iodine/THF gave the predicted 1,3,5-substituted arene 1-(*tert*-butyl)-3,5-diiodobenzene (**9c**) (67% isolated yield).

The DomM and DmmM concepts established here appear influenced by both electronic and steric effects. First, consistent with DoM chemistry, strong DG-containing arenes require shorter reaction times and lower temperatures as compared with those with weak DGs (Table 1, entries 6 to 8 versus entries 1 to 5) to achieve efficient dimetalation. Second, the spatial nature of the DG can redirect the deprotonations toward

**Table 1. Regioselective 2,5-dimagnesiumation of DG-substituted arene substrates and subsequent electrophilic addition reactions.**

| Entry | Substrate | Metalation conditions* | Metalation (%)† | Electrophile                     | Product (yield, %)   |
|-------|-----------|------------------------|-----------------|----------------------------------|--|
| 1     |           | 101°C, 12 hours        | 85              | D <sub>2</sub> O                 | <br><b>6a</b> (80)‡ (91)§  |
| 2     |           | 101°C, 12 hours        | 85              | I <sub>2</sub> / CO <sub>2</sub> | <br><b>6a'</b> (78)  <br><b>6a''</b> (68)                          |
| 3     |           | 101°C, 5 hours         | 88              | I <sub>2</sub> / CO <sub>2</sub> | <br><b>6b'</b> (73)  <br><b>6b''</b> (65)                          |
| 4     |           | 25°C, 12 hours         | 83              | I <sub>2</sub>                   | <br><b>6c'</b> (72)  |
| 5     |           | 101°C, 3.5 hours       | 80              | I <sub>2</sub>                   | <br><b>6d'</b> (67)  |
| 6     |           | 25°C, 3.5 hours        | 91              | I <sub>2</sub> / CO <sub>2</sub> | <br><b>6e'</b> (82)  <br><b>6e''</b> (65)                          |
| 7     |           | 25°C, 1 hour           | 87              | I <sub>2</sub>                   | <br>( <i>P</i> )- <b>6f</b> (33)  <br>( <i>M</i> )- <b>6f</b> (36) |
| 8     |           | 25°C, 20 min           | 91              | I <sub>2</sub>                   | <br><b>6g'</b> (77)  |

\*Metalation conditions (temperature, time).

†Extent of 2,5-dimetalation determined by NMR spectroscopy.

‡Conversion determined by NMR spectroscopy based on in situ reaction of **2<sub>Bu</sub>** with **4a**.

§Isolated yield based on the reaction of crystalline **5a** and D<sub>2</sub>O.

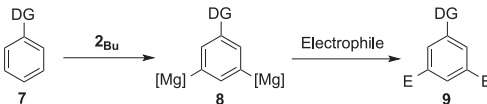
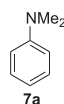
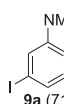
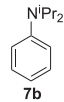
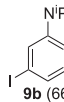
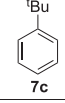
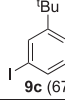
||Isolated yield.

DmmM, breaking the dogma of DoM effects. However, the major driver for dimetalation occurring in the first place seems to be the pre-organized base structure having two butyl ligands on two magnesium atoms primed for executing deprotonation. Further proof of the greater influence of the base over the substrate comes from the observation that the regioselectivity of **2<sub>Bu</sub>** with toluene can be switched from DomM to DmmM by changing the alkyl component from butyl to trimethylsilylmethyl in **2<sub>Me<sub>3</sub>SiCH<sub>2</sub></sub>**

(32). Dimetalation is often seen as an unwanted side reaction complication in organic synthesis, whereas here, it is a deliberate regioselective process controlled by the template nature of the bases  $[\{K[Mg(TMP)_2(^nBu)_6]\}_6]$  **1<sub>Bu</sub>** and  $[\{Na_4Mg_2(TMP)_6(^nBu)_2\}]$  **2<sub>Bu</sub>**. This concept could have a vast scope not just for dimetalation but for metalation in general, provided different template bases can be designed. A glimpse of this scope can be seen in the ring-cage hybrid structure of  $[\{Li(TMP)Li(C_5H_4)_4Li_6(^nBu)_2\}]$  prepared via a shape-selective

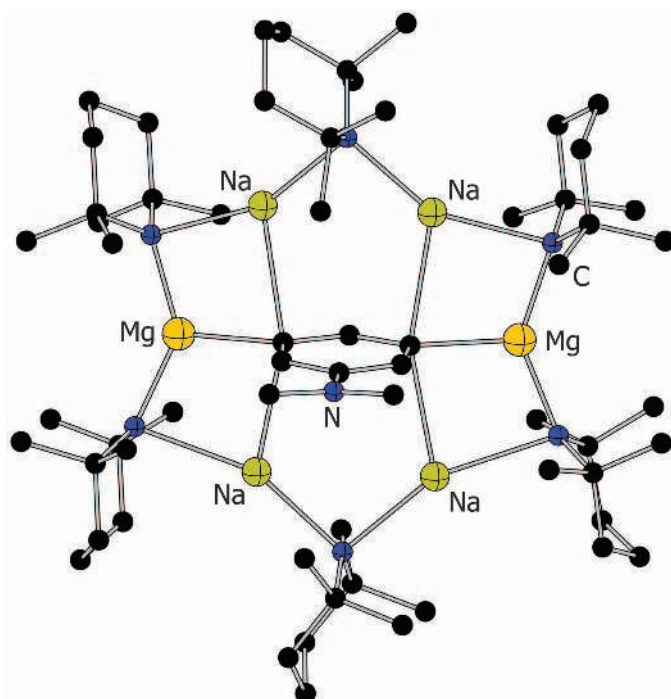
synthesis from the molecular square  $[\{Li(TMP)Li(C_5H_5)_4\}]$  and butyllithium (**33**). This templation approach presently realizes dimetalation reactions exhibiting dual ortho-meta' or meta-meta' regioselectivities, but the principle it establishes bodes well for a broader scope of positional C-H to C-M transformations in the future. Rationally designing a range of template bases is the next challenge for the metalation community.

**Table 2. Regioselective 3,5-dimagnesiation of DG-substituted aryl substrates and subsequent electrophilic addition reaction.**

|  |  |                        |                 |                |   |
|---|--|------------------------|-----------------|----------------|---|
| Entry   | Substrate  | Metalation conditions* | Metalation (%)† | Electrophile   | Product (yield, %)‡   |
| 1   |   | 101°C, 4 hours         | 91              | I <sub>2</sub> |  <b>9a</b> (71)  |
| 2   |   | 101°C, 3 hours         | 85              | I <sub>2</sub> |  <b>9b</b> (66)  |
| 3   |  | 60°C, 12 hours         | 88              | I <sub>2</sub> |  <b>9c</b> (67) |

\*Metalation conditions (temperature, time).  
‡Isolated yield.

†Extension of 3,5-dimetalation determined by NMR spectroscopy.



**Fig. 4. Molecular structure of  $[Na_4Mg_2(TMP)_6(C_6H_3NMe_2-3,5)]$  **8a**.** Hydrogen atoms have been omitted for clarity.

## REFERENCES AND NOTES

- M. R. Hickey et al., *Org. Process Res. Dev.* **9**, 764–767 (2005).
- M. Schlosser, *Organometallics in Synthesis Third Manual*, M. Schlosser, Ed. (John Wiley & Sons, Hoboken, NJ, 2013).
- P. Beak, V. Snieckus, *Acc. Chem. Res.* **15**, 306–312 (1982).
- T. K. Macklin, V. Snieckus, in *Handbook of C-H Transformations*, G. Dyker, Ed. (Wiley-VCH, Weinheim, Germany, 2005), vol. 1, pp. 106–118.
- V. Snieckus, *Chem. Rev.* **90**, 879–933 (1990).
- M. C. Whisler, S. MacNeil, V. Snieckus, P. Beak, *Angew. Chem. Int. Ed.* **43**, 2206–2225 (2004).
- H. Gilman, R. L. Bebb, *J. Am. Chem. Soc.* **61**, 109–112 (1939).
- G. Wittig, U. Pockels, H. Droge, *Ber. Dtsch. Chem. Ges.* **71**, 1903–1912 (1938).
- P. Beak, R. A. Brown, *J. Org. Chem.* **42**, 1823–1824 (1977).
- J. Mortier, J. Moyroud, B. Bennetau, P. A. Cain, *J. Org. Chem.* **59**, 4042–4044 (1994).
- M. Kauch, D. Hoppe, *Synthesis* **2006**, 1578–1589 (2006).
- M. Kauch, D. Hoppe, *Synthesis* **2006**, 1575–1577 (2006).
- T. K. Macklin, V. Snieckus, *Org. Lett.* **7**, 2519–2522 (2005).
- C. Metallinos, S. Nerdinger, V. Snieckus, *Org. Lett.* **1**, 1183–1186 (1999).
- B. Haag, M. Mosrin, H. Ila, V. Malakhov, P. Knochel, *Angew. Chem. Int. Ed.* **50**, 9794–9824 (2011).
- A. Krasovskiy, V. Krasovskaya, P. Knochel, *Angew. Chem. Int. Ed.* **45**, 2958–2961 (2006).
- R.-Y. Tang, G. Li, J.-Q. Yu, *Nature* **507**, 215–220 (2014).
- D. Leow, G. Li, T.-S. Mei, J.-Q. Yu, *Nature* **486**, 518–522 (2012).
- A. J. Martinez-Martinez et al., *Chem. Sci.* **5**, 771–781 (2014).
- A. R. Alborés et al., *Chem. Commun.* **50**, 10588–10591 (2005).
- W. Bauer, D. Seebach, *Helv. Chim. Acta* **67**, 1972–1988 (1984).
- R. E. Mulvey, *Acc. Chem. Res.* **42**, 743–755 (2009).
- P. Alborés et al., *Angew. Chem. Int. Ed.* **48**, 3317–3321 (2009).
- C. Schade, W. Bauer, P. Von Ragué Schleyer, *J. Organomet. Chem.* **295**, C25–C28 (1985).
- Materials and methods are available as supplementary materials on Science Online.
- K. Flick, E. Frankus, U.S. patent US3652589A (1972).
- M. Protiva, *Drugs Future* **16**, 911–916 (1991).
- W. Bauer, P. V. Schleyer, *J. Am. Chem. Soc.* **111**, 7191–7198 (1989).
- S. T. Chadwick, R. A. Rennels, J. L. Rutherford, D. B. Collum, *J. Am. Chem. Soc.* **122**, 8640–8647 (2000).
- S. Harder, J. Boersma, L. Brandsma, J. A. Kanters, *J. Organomet. Chem.* **339**, 7–15 (1988).
- J. M. Saa, P. M. Deya, G. A. Suer, A. Frontera, *J. Am. Chem. Soc.* **114**, 9093–9100 (1992).
- V. L. Blair et al., *Angew. Chem. Int. Ed.* **47**, 6208–6211 (2008).
- A. A. Fyfe, A. R. Kennedy, J. Klett, R. E. Mulvey, *Angew. Chem. Int. Ed.* **50**, 7776–7780 (2011).

## ACKNOWLEDGMENTS

We gratefully acknowledge financial support from the Royal Society (Wolfson research merit award to R.E.M.) and the Engineering and Physical Sciences Research Council (EPSRC) (Career Acceleration Fellowship, EP/J001872/1 and EP/L001497/1 to C.T.O.H.). E. Hevia and B. J. Fleming are also thanked for insightful discussions. We are also grateful to the EPSRC UK National Mass Spectrometry Facility (NMSF), Swansea for collecting our mass spectral data. Metrical data for **5a** and **8a** are freely available from the Cambridge Crystallographic Database Centre (CCDC-1017351 and CCDC-1017352, respectively).

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6211/834/suppl/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S80  
Tables S1 to S3  
References (34–51)

5 August 2014; accepted 3 October 2014  
10.1126/science.1259662



## EXOPLANET ATMOSPHERE

# Thermal structure of an exoplanet atmosphere from phase-resolved emission spectroscopy

Kevin B. Stevenson,<sup>1,13\*</sup> Jean-Michel Désert,<sup>2</sup> Michael R. Line,<sup>3</sup> Jacob L. Bean,<sup>1</sup> Jonathan J. Fortney,<sup>3</sup> Adam P. Showman,<sup>4</sup> Tiffany Kataria,<sup>4</sup> Laura Kreidberg,<sup>1</sup> Peter R. McCullough,<sup>5,6</sup> Gregory W. Henry,<sup>7</sup> David Charbonneau,<sup>8</sup> Adam Burrows,<sup>9</sup> Sara Seager,<sup>10</sup> Nikku Madhusudhan,<sup>11</sup> Michael H. Williamson,<sup>7</sup> Derek Homeier<sup>12</sup>

Exoplanets that orbit close to their host stars are much more highly irradiated than their solar system counterparts. Understanding the thermal structures and appearances of these planets requires investigating how their atmospheres respond to such extreme stellar forcing. We present spectroscopic thermal emission measurements as a function of orbital phase ("phase-curve observations") for the highly irradiated exoplanet WASP-43b spanning three full planet rotations using the Hubble Space Telescope. With these data, we construct a map of the planet's atmospheric thermal structure, from which we find large day-night temperature variations at all measured altitudes and a monotonically decreasing temperature with pressure at all longitudes. We also derive a Bond albedo of  $0.18_{-0.07}^{+0.07}$  and an altitude dependence in the hot-spot offset relative to the substellar point.

Previous exoplanet phase-curve observations (1–7) have revealed day-night temperature contrasts and hot-spot offsets relative to the substellar point (the point at which the host star would be perceived to be directly overhead). However, these observations were limited to broadband photometry; therefore, the altitudes probed by the phase curves were not uniquely constrained. Spectroscopic phase curves can break previous degeneracies by permitting us to uniquely identify the main atmospheric opacity

source within the observed bandpass and infer the planet's atmospheric temperature-pressure profile as a function of orbital phase (8–12).

The WASP-43 system contains a transiting Jupiter-size exoplanet on a 19.5-hour orbit around its K4 host star (13, 14). Previous measurements (14–17) of its dayside thermal emission detected no signs of a thermal inversion and suggested low day-night energy redistribution. However, the precise thermal structure of the dayside atmosphere remains unknown without higher-resolution ob-

servations, and the planet's global energy budget and atmospheric heat-redistribution efficiency are poorly constrained without observations of the nightside.

Over 4 to 7 November 2013, we used the Wide Field Camera 3 (WFC3) instrument from the Hubble Space Telescope (HST) to observe three nearly consecutive orbits of WASP-43b. The planet orbits so close to its host star that it is tidally locked. Therefore, orbital phase is equivalent to rotational phase for the planet, and observations over a complete orbit allow us to map the entire surface of the planet. HST also acquired data for three primary transits and two secondary eclipses, where the planet passes in front of and behind its host star, respectively, between 9 November 2013 and 5 December 2013.

<sup>1</sup>Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA.

<sup>2</sup>CASA, Department of Astrophysical and Planetary Sciences, University of Colorado, 389-UCB, Boulder, CO 80309, USA.

<sup>3</sup>Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA.

<sup>4</sup>Department of Planetary Sciences and Lunar and Planetary Laboratory, The University of Arizona, Tucson, AZ 85721, USA.

<sup>5</sup>Space Telescope Science Institute, Baltimore, MD 21218, USA.

<sup>6</sup>Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA.

<sup>7</sup>Center for Excellence in Information Systems, Tennessee State University, Nashville, TN 37209, USA.

<sup>8</sup>Department of Astronomy, Harvard University, Cambridge, MA 02138, USA.

<sup>9</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA.

<sup>10</sup>Department of Earth, Atmospheric, and Planetary Sciences, Department of Physics, Massachusetts Institute of Technology, 54-1718, 77 Massachusetts Avenue, Cambridge, MA 02139, USA.

<sup>11</sup>Institute of Astronomy, University of Cambridge, Cambridge CB3 0HA, UK.

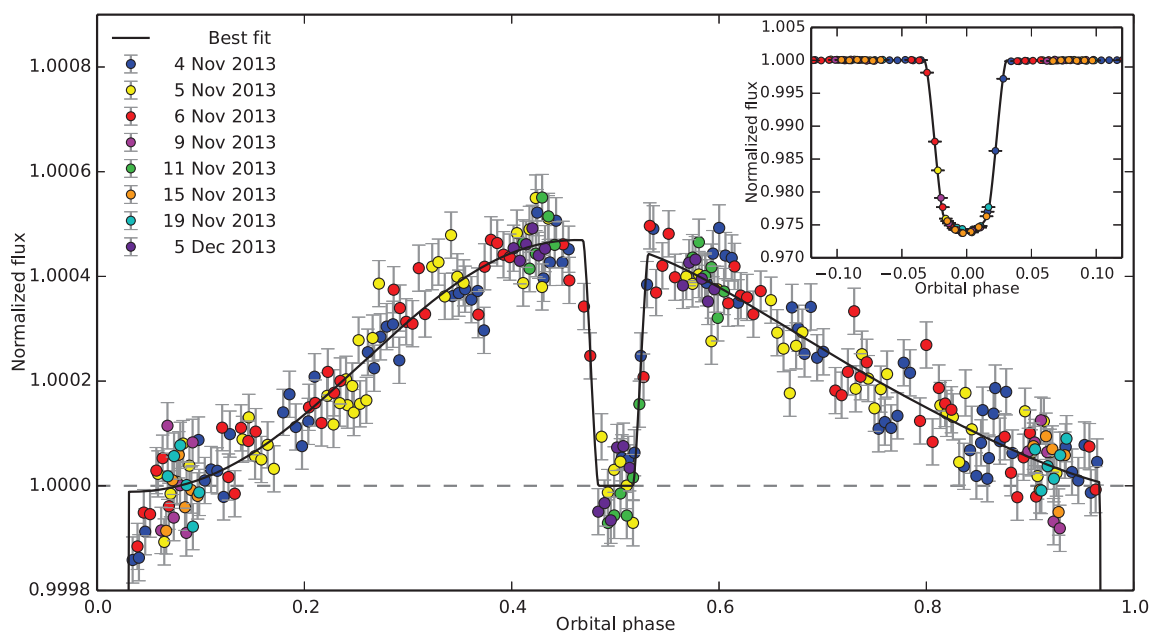
<sup>12</sup>Centre de Recherche Astrophysique de Lyon, UMR 5574, CNRS, Université de Lyon, École Normale Supérieure de Lyon, 46 Allée d'Italie, F-69364 Lyon Cedex 07, France.

<sup>13</sup>NASA Sagan Fellow.

\*Corresponding author. E-mail: kbs@uchicago.edu

**Fig. 1. Band-integrated phase curve of WASP-43b.**

The systematics-corrected flux values are binned in time, normalized to the stellar flux, and have 1 $\sigma$  error bars. Each color represents data acquired from a different HST visit. The phase curve depicts steadily increasing and decreasing observed flux that originates from different longitudes of the tidally locked planet as it makes one complete rotation. Light from the planet is blocked near an orbital phase of 0.5 as it is eclipsed by its host star. The model phase curve maximum occurs  $40 \pm$



3 min before the midpoint of secondary eclipse, which corresponds to a shift of  $12.3 \pm 1.0^\circ$  east of the substellar point. The model phase curve minimum occurs  $34 \pm 5$  min after the primary transit midpoint, or  $10.6 \pm 1.4^\circ$  west of the antistellar point. As a result, maximum planetary emission occurs  $0.436 \pm 0.005$  orbits after the observed minimum (for depths probed by these observations) and the shape of the phase curve is asymmetric. Inset, for comparison, is the white light curve primary transit. It is notable that the observed flux values are consistently low for  $\sim 30$  min after transit egress.

All of the observations used the G141 grism (1.1 to 1.7  $\mu\text{m}$ ) and the bidirectional spatial scan mode.

Using custom software (18, 19), we reduced the data and extracted the spectra. We produced time-series spectroscopy by dividing the spectra into 15 0.035- $\mu\text{m}$ -wide channels (7 pixels, resolution  $R = \lambda/\Delta\lambda \sim 37$ ). We also produced band-integrated “white” light curves to resolve finer details in the shape of the phase curve (Fig. 1). We simultaneously fit the light curves using transit and uniform-source eclipse models (20), a baseline flux for each HST scan direction, two standard model components for HST orbit-long and visit-long systematics, and a sinusoidal function to represent the phase variation (19, 21). We estimate uncertainties using a differential-evolution Markov-chain Monte Carlo (DE-MCMC) algorithm (18) and utilize an independent analysis pipeline (21) to confirm our light-curve fits.

The white light phase curve (Fig. 1) reveals a distinct increase in flux as the tidally locked day-side rotates into view. The flux peaks prior to secondary eclipse (eastward of the substellar point) and then decreases until the planet tran-

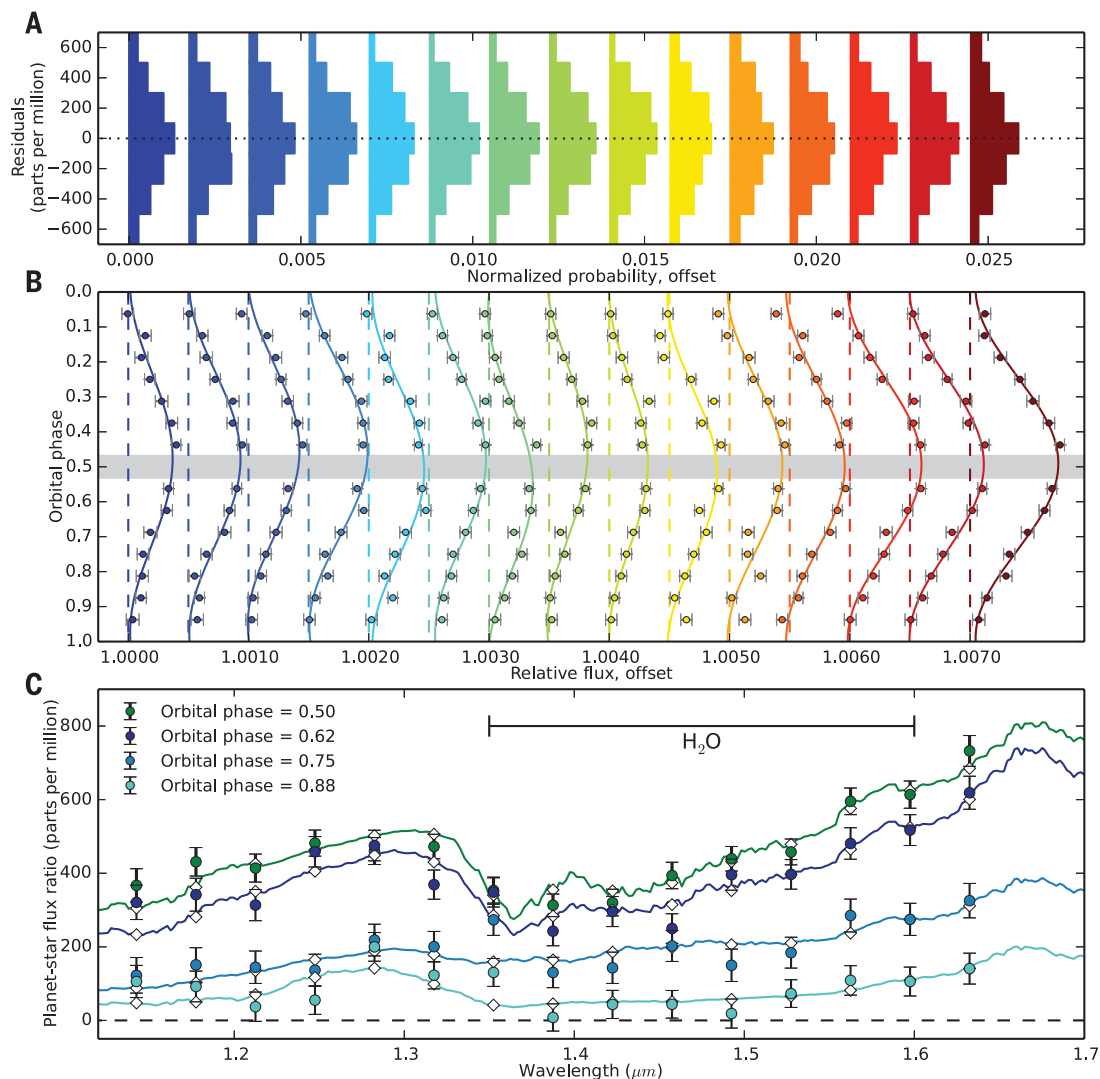
sits in front of its host star. Because the phase curve minimum occurs west of the antistellar point, we detect a strong asymmetry ( $\sim 10\sigma$ ) in the shape of the observed phase curve. We measure a white light curve eclipse depth that is consistent with the peak-to-peak planet flux variation. This confirms a relatively cool night side and poor heat redistribution. Table S1 lists our best-fit parameters with uncertainties.

We gain additional information by decomposing the white light phase curve into 15 spectrophotometric channels (Fig. 2). The spectrally resolved phase curves exhibit wavelength-dependent amplitudes, phase shifts, and eclipse depths (table S2). We use the measured phase-resolved emission spectra (Fig. 2C) to infer the temperature structure and molecular abundances at 15 binned orbital phases (each of width 0.0625). We fit atmospheric models to these spectra using a DE-MCMC approach from the CHIMERA Bayesian retrieval suite (22). For each phase, a five-parameter, double-gray radiative equilibrium solution parameterizes the planet's temperature structure (23). The models include six thermochemically

plausible and spectrally prominent absorbers ( $\text{H}_2\text{O}$ ,  $\text{CH}_4$ ,  $\text{CO}$ ,  $\text{CO}_2$ ,  $\text{NH}_3$ , and  $\text{H}_2\text{S}$ ). We find that water is the only absorber to significantly influence the phase-resolved emission spectra (24) (Fig. 2). The model spectra are in good agreement with the data, achieving a typical  $\chi^2$  value of 18 with 15 data points and 6 relevant free parameters (fig. S3).

Using the atmospheric models to estimate the day- and night-side fluxes, we find that the planet redistributes heat poorly [(19), heat redistribution factor  $\mathcal{F} = 0.503^{+0.021}_{-0.003}$ , where  $\mathcal{F} = 0.5 \rightarrow 1$  spans the range from zero to full heat redistribution]. This is predicted to occur when the radiative time scale is shorter than the relevant dynamical time scales, including those for wave propagation and advection over a hemisphere (25). Poor redistribution has been inferred before, but only for hot Jupiters receiving significantly greater stellar flux than WASP-43b (4, 7). We estimate the fraction of incident stellar light reflected by WASP-43b's atmosphere by computing the day- and night-side bolometric fluxes from the model spectra and find a Bond albedo of  $0.18^{+0.07}_{-0.12}$ . This

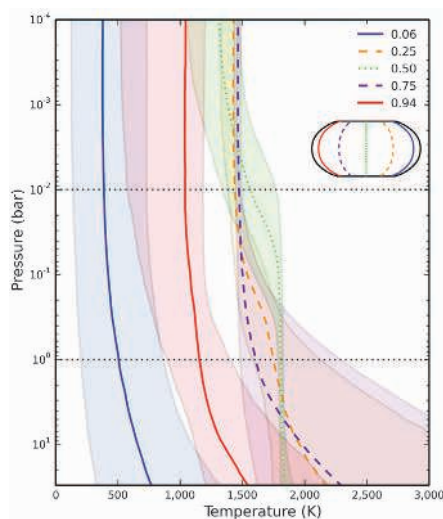
**Fig. 2. Phase-resolved emission spectrum of WASP-43b relative to the stellar flux.** (A) The histograms of the unbinned phase-curve residuals are separated horizontally by wavelength [colors, defined on the abscissa of (C)] for clarity. The residuals are Gaussian distributed with a zero mean and show no evidence of correlated noise. (B) We show binned phase curves (colored points with  $1\sigma$  error bars) and best-fit models (colored lines). The planet emission is normalized with respect to the stellar flux and separated horizontally by wavelength for clarity. The gray region depicts the time of secondary eclipse. (C) We illustrate a subset of data points from (B), except plotted as a function of wavelength and with best-fit atmospheric models (colored lines). White diamonds depict the models binned to the resolution of the data. For clarity, we only display planet-to-star flux ratios at four planet phases: full (0.5, secondary eclipse), waning gibbous (0.62), half (0.75), and waning crescent (0.88). In figs. S1 to S3, we provide full 1D and 2D representations of (B) and (C). A time-lapse video of the planet's phase-resolved emission spectrum is available in movie S1.





method assumes energy balance with the parent star but requires no detection of reflected light (19). The low Bond albedo is consistent with model predictions that hot Jupiters absorb most of the flux incident upon them (11, 26, 27).

The atmospheric model fits reveal information about WASP-43b's phase-dependent thermal structure at the pressure levels probed by these observations (Fig. 3). Depending on the wavelength and phase, these pressures range from 0.01 to 1 bar (fig. S4). The retrieved thermal profiles are consistent with a global, monotonically decreasing temperature with altitude, as would be expected from radiative cooling without high-altitude absorbers of stellar radiation. As a test, we compare the retrieved dayside-averaged thermal profile to three scenarios of self-consistent radiative equilibrium models (28) and find that it is most congruous with the



**Fig. 3. Thermal profiles of WASP-43b at select orbital phases.** Higher pressures indicate deeper within the planet's atmosphere. Colored curves depict median values with  $1\sigma$  uncertainty regions for the assumed parameterization of the retrieval. We illustrate the temperature asymmetry on the planet's night side immediately before and after transit (orbital phase = 0.06 and 0.94), the similar thermal profiles on WASP-43b's morning and evening terminators (0.25 and 0.75), and the dayside-averaged profile (0.5). The HST/WFC3 measurements probe the atmosphere primarily between 0.01 and 1.0 bar (horizontal dotted lines). The retrieved model profiles are 1D representations of the disk-integrated flux values at each phase. However, because the emitted flux values at these wavelengths are near the peak of the Planck curve, the flux scales as  $T^5$  or more and the disk-integrated thermal profiles are heavily weighted toward the hotter dayside. As a result, there is no significant change in the modeled temperature structure over half of the orbital phases (0.25  $\rightarrow$  0.75, when the substellar point is visible). We plot individual pressure-temperature profiles with  $1\sigma$  uncertainty regions in fig. S4. A time-lapse video of WASP-43b's phase-resolved thermal profile is available in movie S1.

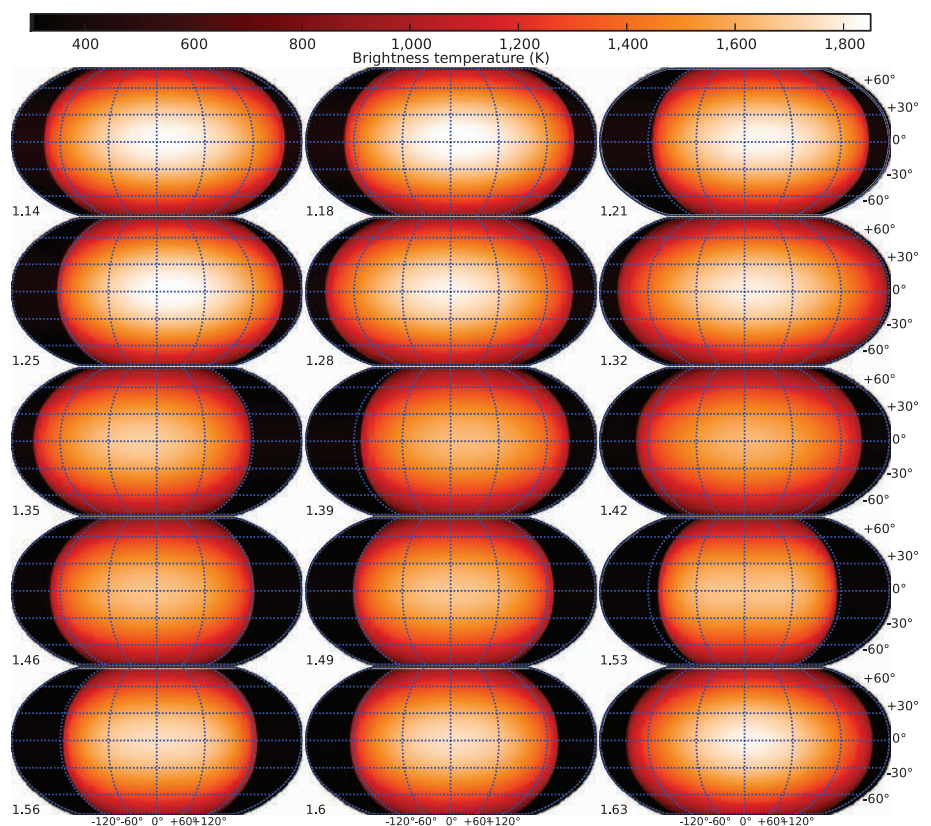
thermal structure expected at the substellar point (fig. S5). This result supports our findings of a low day-night heat redistribution.

Adopting the same sinusoidal function used to fit the phase variation (19), we invert the spectroscopic light curves into longitudinally resolved brightness temperature maps (29) (Fig. 4). The brightness temperature,  $T_B$ , is a function of atmospheric opacity, and water vapor is the main source of opacity in this bandpass. Because  $T_B$  is systematically cooler within the water band, this signifies the global presence of water vapor within the pressure regions probed by these measurements (fig. S7).

The large measured day-night luminosity difference of WASP-43b [ $(19), L_{\text{day}}/L_{\text{night}} > 20$  at 1 $\sigma$ , mode  $\sim 40$ ] contrasts sharply with the modest day-night differences inferred from Spitzer Space Telescope photometry for giant planets such as HD 189733b, HD 209458b, and HD 149026b that are similarly irradiated (1, 5, 25). Unlike Spitzer

data, our spectrum samples the planet's flux near the peak of its Planck curve, allowing for a more robust determination of the total dayside luminosity. This data set suggests that derived day-night differences may be strongly wavelength dependent and that mid-infrared photometry may not give a complete picture of planetary circulation.

Brightness temperature maps, being functions of both longitude and atmospheric depth, reveal the dynamics of a planet's atmosphere. Phase-curve peaks prior to the time of secondary eclipse (as seen in Fig. 1) have previously been reported in hot Jupiters (1, 6) and match predictions from three-dimensional (3D) circulation models (30–32). Such models show that the eastward offset results from a strong jet stream at the equator; our observations thus suggest that WASP-43b exhibits such an eastward-flowing jet. Our spectrophotometric observations further demonstrate the influence of water vapor on the emergent thermal structure. Inside the water band (1.35 to



**Fig. 4. Longitudinally resolved brightness temperature maps of WASP-43b in all 15 spectro-photometric channels.** Black regions in this Robinson projection indicate no discernible contribution. Numbers indicate the wavelength (in micrometers). The observations constrain the brightness temperature at each longitude, but contain no latitudinal information (we assign a  $\cos^2$  weighting). In general, the change in temperature is relatively small over the planet's dayside ( $-90^\circ$  to  $+90^\circ$ ) and comparatively extreme near  $\pm 120^\circ$ , thus indicating that we detect emission over the planet's entire dayside. Because WASP-43b does not contain a thermal inversion at these pressures, the hotter regions at a given longitude sample deeper within the atmosphere. The presence of water vapor in the planet's atmosphere explains the relatively cool brightness temperature from 1.35 to 1.6  $\mu\text{m}$ . Outside of the water feature, the brightness temperature peak (indicated in white) is predominantly eastward (toward positive longitudes) of the substellar point. This correlation is readily seen in fig. S8 and matches the predictions of 3D circulation models. Figure S6 displays 1D brightness temperatures with uncertainty regions.

1.6  $\mu\text{m}$ ), observations probe lower atmospheric pressures (higher altitudes), and we measure smaller phase-curve peak offsets relative to the other wavelengths (figs. S7 and S8). This is qualitatively consistent with variable brown dwarf measurements (33) and circulation-model predictions (25, 31, 32, 34), which show that smaller displacements are expected at higher altitudes where radiative time scales are much shorter than the relevant dynamical time scales. However, the observed westward offset of the coldest regions from the antistellar point is puzzling and is not predicted by most models.

The strong day-night temperature variation observed for WASP-43b distinguishes itself from the predominantly uniform temperatures of the solar system giant planets. This illustrates the importance of radiative forcing on the atmospheres of close-in exoplanets. Phase-resolved emission spectroscopy offers a unique way to determine how the extreme stellar radiation incident on these planets is absorbed, circulated, and reemitted. This approach represents a new opportunity for future observations to constrain theories of planetary atmospheric dynamics in a new regime.

#### REFERENCES AND NOTES

- H. A. Knutson *et al.*, *Nature* **447**, 183–186 (2007).
- H. A. Knutson *et al.*, *Astrophys. J.* **703**, 769–784 (2009).
- I. J. M. Crossfield *et al.*, *Astrophys. J.* **723**, 1436–1446 (2010).
- N. B. Cowan *et al.*, *Astrophys. J.* **747**, 82 (2012).
- H. A. Knutson *et al.*, *Astrophys. J.* **754**, 22 (2012).
- N. K. Lewis *et al.*, *Astrophys. J.* **766**, 95 (2013).
- P. F. L. Maxted *et al.*, *Mon. Not. R. Astron. Soc.* **428**, 2645–2660 (2013).
- S. Seager, D. D. Sasselov, *Astrophys. J.* **502**, L157–L161 (1998).
- A. Burrows *et al.*, *Astrophys. J.* **534**, L97–L100 (2000).
- T. Guillot, A. P. Showman, *Astron. Astrophys.* **385**, 156–165 (2002).
- D. Sudarsky, A. Burrows, I. Hubeny, *Astrophys. J.* **588**, 1121–1148 (2003).
- J. J. Fortney, C. S. Cooper, A. P. Showman, M. S. Marley, R. S. Freedman, *Astrophys. J.* **652**, 746–757 (2006).
- C. Hellier *et al.*, *Astron. Astrophys.* **535**, L7 (2011).
- M. Gillon *et al.*, *Astron. Astrophys.* **542**, A4 (2012).
- W. Wang *et al.*, *Astrophys. J.* **770**, 70 (2013).
- G. Chen *et al.*, *Astron. Astrophys.* **563**, A40 (2014).
- J. Blecic *et al.*, *Astrophys. J.* **781**, 116 (2014).
- K. B. Stevenson *et al.*, *Astron. J.* **147**, 161 (2014).
- Materials and methods are available as supplementary materials on Science Online.
- K. Mandel, E. Agol, *Astrophys. J.* **580**, L171–L175 (2002).
- L. Kreidberg *et al.*, *Nature* **505**, 69–72 (2014).
- M. R. Line, H. Knutson, A. S. Wolf, Y. L. Yung, *Astrophys. J.* **783**, 70 (2014).
- V. Parmentier, G. Guillot, *VizieR Online Data Catalog* **356**, 29133 (2013).
- L. Kreidberg *et al.*, *Astrophys. J.* **793**, L27 (2014).
- D. Perez-Becker, A. P. Showman, *Astrophys. J.* **776**, 134 (2013).
- M. S. Marley, C. Gelino, D. Stephens, J. I. Lunine, R. Freedman, *Astrophys. J.* **513**, 879–893 (1999).
- A. Burrows, L. Ibgui, I. Hubeny, *Astrophys. J.* **682**, 1277–1282 (2008).
- J. J. Fortney, K. Lodders, M. S. Marley, R. S. Freedman, *Astrophys. J.* **678**, 1419–1435 (2008).
- N. B. Cowan, E. Agol, *Astrophys. J.* **678**, L129–L132 (2008).
- A. P. Showman, T. Guillot, *Astron. Astrophys.* **385**, 166–180 (2002).
- C. S. Cooper, A. P. Showman, *Astrophys. J.* **629**, L45–L48 (2005).
- A. P. Showman *et al.*, *Astrophys. J.* **699**, 564–584 (2009).
- E. Buenzli *et al.*, *Astrophys. J.* **760**, L31 (2012).
- A. Burrows, E. Rauscher, D. S. Spiegel, K. Menou, *Astrophys. J.* **719**, 341–350 (2010).

#### ACKNOWLEDGMENTS

This work is based on observations made with the NASA–European Space Agency Hubble Space Telescope that were obtained at the Space Telescope Science Institute (STScI), which is operated

by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. Data are available through the Mikulski Archive for Space Telescopes (MAST). We thank A. Vick and M. Reinhard of STScI for scheduling these observations, which are associated with program GO-13467. Support for this work was provided by NASA through a grant from the STScI, the Sagan Fellowship Program (to K.B.S.) as supported by NASA and administered by the NASA Exoplanet Science Institute (NExSci), the Alfred P. Sloan Foundation through a Sloan Research Fellowship (to J.L.B.), and the NSF through a Graduate Research Fellowship (to L.K.). G.W.H. and M.H.W. acknowledge long-term support from NASA, NSF, Tennessee State University, and the State of Tennessee through its Centers of Excellence program. S.S. acknowledges funding from the Massachusetts Institute of Technology. D.H. acknowledges support from the European

Research Council under the European Community's Seventh Framework Programme—FP7/2007–2013 grant agreement no. 247060.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6211/838/suppl/DC1  
Materials and Methods  
Supplementary Text  
Tables S1 and S2  
Figs. S1 to S8  
Movie S1

30 May 2014; accepted 17 September 2014  
Published online 9 October 2014;  
10.1126/science.1256758

#### SEXUAL CONFLICT

## The evolution of infanticide by males in mammalian societies

Dieter Lukas<sup>1\*</sup> and Elise Huchard<sup>1,2</sup>

Male mammals often kill conspecific offspring. The benefits of such infanticide to males, and its costs to females, probably vary across mammalian social and mating systems. We used comparative analyses to show that infanticide primarily evolves in social mammals in which reproduction is monopolized by a minority of males. It has not promoted social counterstrategies such as female gregariousness, pair living, or changes in group size and sex ratio, but is successfully prevented by female sexual promiscuity, a paternity dilution strategy. These findings indicate that infanticide is a consequence, rather than a cause, of contrasts in mammalian social systems affecting the intensity of sexual conflict.

Infanticide by males is widespread in mammals and may be the main cause of infant mortality in some populations (1). It has long been viewed as a sexually selected strategy that increases mating opportunities for killer males by shortening postpartum infertility in the victim's mother (2, 3). This is supported by natural observations across taxa showing that males target unrelated infants and often impregnate the victim's mother afterward (1, 3, 4). Attempts to explain the taxonomic distribution of infanticide have, however, been mainly limited to investigations of the life-history correlates of male infanticide (5–8), showing that it rarely occurs where it does not accelerate the mother's return to sexual activity, as in seasonal breeders who cannot resume cycling before the next breeding season (8).

Several scenarios have linked the distribution of infanticide across species to the evolution of social organization and mating systems in mammals, and they remain largely untested or disputed (9, 10). According to the sexual selection hypothesis, the distribution of infanticide is ex-

pected to be modulated by contrasts in social systems that affect the intensity of male intra-sexual competition (3, 5, 8). Male infanticide should thus be prevalent in species in which a few males mate with multiple females and monopolize most reproductive opportunities. However, females may respond to infanticide by developing counterstrategies that may refine patterns of associations across species. Social counterstrategies may include the evolution of female sociality (6, 11, 12), of permanent male-female associations (13–15), or of changes in the group sex ratio (6, 11, 12, 15), because females may form coalitions with other females or with resident males to defend their progeny against male invaders. Additionally, females may mate with multiple partners to confuse paternity and dissuade infanticide (6, 16, 17), which may thus be absent in species with pronounced sperm competition. The evolutionary arms race between the sexes driven by male infanticide has generated confusion and controversy regarding the role of infanticide in the evolution of mammalian societies, calling for more integrative studies.

We used information gathered on 260 mammal species, including 119 species with and 141 without infanticide (Fig. 1), to perform phylogenetic analyses (18) to identify how variation in social organization and mating systems may have favored or prevented the evolution of infanticide by males. Specifically, we tested whether

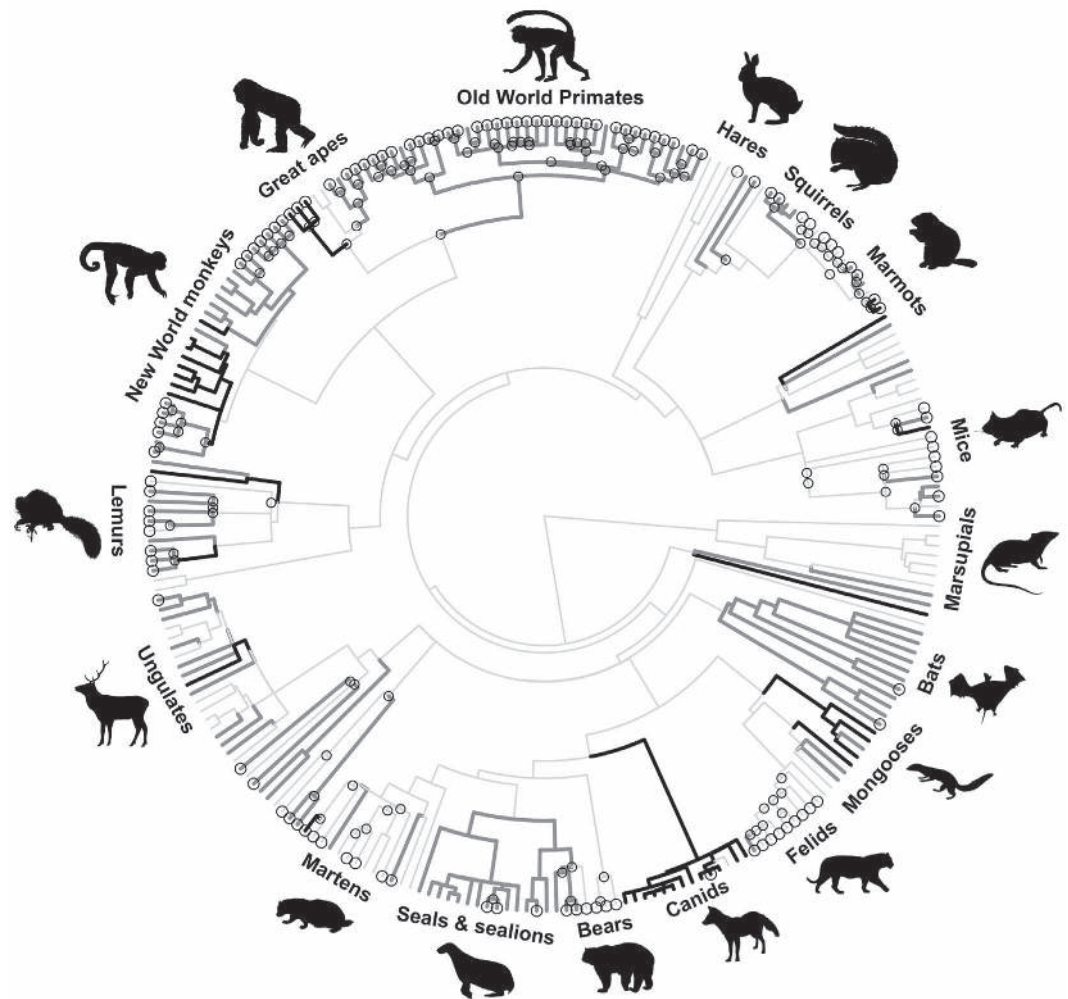
<sup>1</sup>Large Animal Research Group, Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK. <sup>2</sup>Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, CNRS – Université de Montpellier, 1919 Route de Mende, 34293 Montpellier Cedex 5, France.

\*Corresponding author. E-mail: dl384@cam.ac.uk



**Fig. 1. Distribution of male infanticide across mammals.**

Male infanticide has been reported for about half of all species in our sample (open circles) and seems to have evolved independently multiple times. It mostly occurs in social species (dark gray branches), less in solitary species (light gray branches), and least in monogamous species (black branches). Animal drawings are from phylopic.org [for full credit, see (18)]



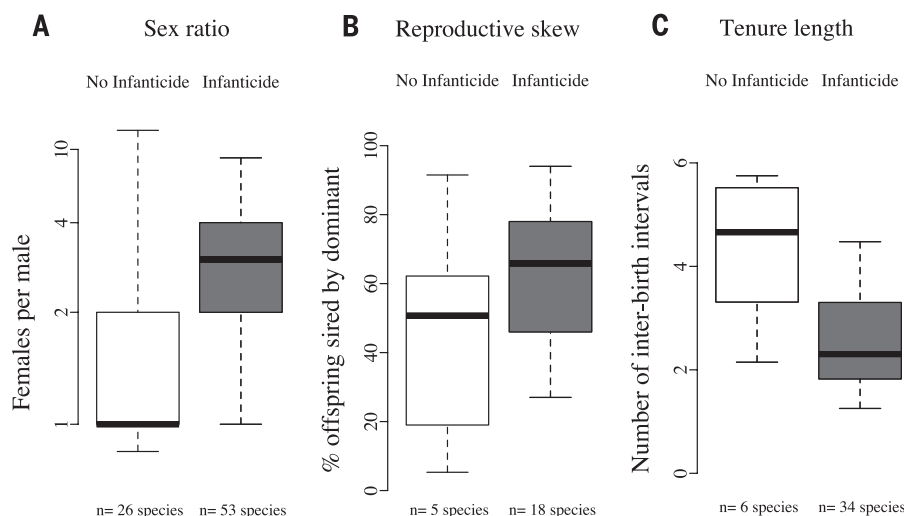
infanticide has primarily evolved in species in which females breed throughout the year and some males have high reproductive monopoly, and whether it has been selected for social and sexual counterstrategies, including transitions toward female gregariousness, permanent male-female associations, changes in group size or sex ratio, or increases in female sexual promiscuity. We included infanticide records from wild populations in which the killer was unambiguously identified as an adult male. Species in which infanticide had never been observed were included only if natural observations on females and juveniles were available over more than three reproductive seasons, to minimize the risk of misclassifying them as “non-infanticidal.” The phylogenetic relatedness between species was inferred from an updated mammalian super-tree (19), and analyses were also replicated in primates exclusively, which confirmed results obtained across mammals (see the supplementary materials).

The distribution of male infanticide is closely associated with the absence of seasonal or annual breeding (defined as an average interbirth interval of 360 to 370 days). Infanticide occurs in a majority of non-annual breeders (76% of 97

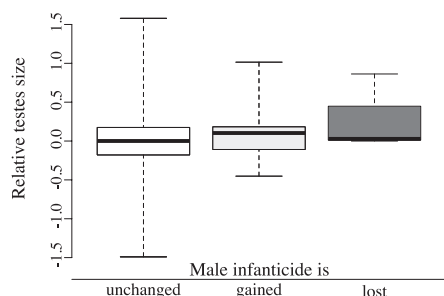
species), as compared to 28% of 134 annual breeders ( $\lambda = 0.67$ ,  $t = -4.0$ ,  $P < 0.001$ ). A long lactation (relative to gestation) and the absence of a postpartum oestrus, two life-history traits affecting the time to cycling resumption in mothers of killed infants, are also associated with male infanticide, although their association disappears when controlling for annual breeding [for both traits,  $P$  as estimated by Markov chain Monte Carlo methods for generalized linear mixed models (pMCMC)  $> 0.08$ , whereas pMCMC for annual breeding  $< 0.001$ ]. Although previous studies have used the ratio of the duration of lactation to that of gestation to estimate infanticide risk (6–8), our results reveal that this association is weak: Lactation exceeds gestation in 25 of 45 mammal species with infanticide and in 38 of 89 species without infanticide. Other aspects of life history were not associated with male infanticide (longevity:  $n = 210$  species, pMCMC = 0.08; litter size:  $n = 230$  species, pMCMC = 0.07; relative offspring weight at birth:  $n = 191$  species, pMCMC = 0.59), suggesting that the evolution of infanticide is not determined by a species’ pace of life. Overall, the possibility for females to breed throughout the year is the only life-history trait examined that significantly explains the distribution of in-

fanticide, confirming previous work and suggesting that infanticidal males gain extra mating opportunities due to earlier cycle resumption of the victim’s mother in non-annual breeders. Such males are thereby redirecting maternal investment from unrelated offspring toward their own future offspring.

Comparing the occurrence of male infanticide across social systems, we find that males commit infanticide more frequently in stable mixed-sex groups (66% of 112 species) than in solitary species [40% of 78 solitary species,  $P$  as estimated by phylogenetic analysis of variance (pAN) = 0.01], in species with female-only groups (23% of 31 species, pAN = 0.006), and in pair-living species (18% of 39 species, pAN  $< 0.001$ ). Phylogenetic reconstructions confirm that variation in social organization is a key determinant of the evolution of male infanticide, which is more likely in stable bisexual groups than in any other social system (table S1). In addition, social groups contain, on average, 1 male per 2.5 females in species with male infanticide ( $SD = 1.7$ ,  $n = 56$  species) but only 1 male for 1.3 females in species without infanticide ( $SD = 3.2$ ,  $n = 19$  species, Fig. 2A;  $\lambda = 0.47$ ,  $t = -3.3$ ,  $P = 0.002$ ; this is not an effect of differences in group sizes, because group



**Fig. 2. Infanticide is associated with polygynous mating systems.** Male infanticide occurs in species in which (A) social groups contain more females per male, (B) dominant males obtain a higher share of reproduction in a given season, but (C) maintain their dominant position for shorter periods (here measured as the average number of interbirth intervals).



**Fig. 3. Loss of male infanticide occurs in species with large testes.** In some lineages in which males commit infanticide, testes appear to increase in size (relative to body mass) and subsequently, male infanticide is lost when testes have become relatively large.

size is similar in both sets of species:  $P = 0.28$ ). Phylogenetic reconstructions suggest that a biased group sex ratio is a determinant, rather than a consequence, of the evolution of male infanticide: The strongest female biases are observed just before or at the transition toward male infanticide, whereas there are few or no subsequent changes once infanticide has evolved (see the supplementary materials).

In line with this, infanticide occurs more frequently where reproduction is monopolized by a few males (Fig. 2B): The median percentage of offspring sired by the alpha male in a group is 67% (SD = 19.9,  $n = 15$  species) in species with infanticide, compared to 35% (SD = 22.2,  $n = 5$ ) in species without (lambda = 0.75,  $t = 3.9$ ,  $P = 0.001$ ). Finally, males retain their dominant position for shorter durations in species with infanticide (Fig. 2C): On average, dominance tenure covers two interbirth intervals (if the females were to successfully wean each litter) (SD = 1.0,  $n = 34$  species), compared to four in species

without infanticide (SD = 1.2,  $n = 6$ ) (lambda = 0.0,  $t = -3.4$ ,  $P = 0.001$ ). These correlations hold after controlling for the effect of annual breeding, which may also affect male ability to monopolize females (table S2). Overall, these findings indicate that a high male monopolization potential favors transitions toward infanticide by allowing males to maximize their reproductive output in systems where the costs of holding dominance shorten their breeding tenures (20).

Male infanticide has been proposed to alter social evolution by favoring female gregariousness or permanent male-female associations or by promoting a higher number of resident males that may help mothers to protect their infants from male intruders. We found no support for any of these scenarios. Infanticide risk is low in solitary species, the ancestral mammalian social organization (9), and as such is unlikely to motivate evolutionary transitions from a solitary lifestyle to other social organizations. Phylogenetic reconstructions confirm that male infanticide is significantly more likely to evolve after transitions toward group living than in their solitary ancestors. The presence of male infanticide does not increase the likelihood of a transition from solitary living to social monogamy, to female gregariousness, or to group living (table S1). In addition, among species living in stable bisexual groups, male infanticide is not linked to changes in the degree of female sociality ( $5.5 \pm 5.1$  females per group in 56 infanticidal species as compared to  $8.9 \pm 4.0$  in 13 non-infanticidal species; pMCMC = 0.93). These findings indicate that infanticide by males did not trigger important transitions in mammalian social evolution. Population studies have nevertheless highlighted relationships between infanticide rates and group size (21), suggesting that the social consequences of infanticide may differ within populations versus across species, or that selection exerted by infanticide is in-

sufficient at inducing major shifts in mammalian social organization.

Infanticide risk may also lead to the evolution of sexual counterstrategies by females, who might prevent infanticide by mating with many males to dilute paternity (6, 16, 17). To test this hypothesis, we investigated evolutionary relationships between infanticide and testis size, an indicator of the intensity of sperm competition (22, 23). In support of the paternity dilution hypothesis, phylogenetic reconstructions indicate that relative testis size is larger in lineages in which more time has elapsed since the evolution of infanticide. State-dependent diversification models (24) show that increases in testis size follow rather than precede transitions to infanticide, because they are not detected on branches immediately before transitions to infanticide, suggesting that infanticide exerts directional selection on relative testis size (see the supplementary materials). In addition, infanticide is subsequently lost in lineages where testes have grown large (Fig. 3; pMCMC = 0.02; see the supplementary results), indicating that infanticide may disappear where female attempts to prevent male sexual monopolization are successful.

Our findings support key predictions of the sexual selection hypothesis: Males primarily kill the offspring of rivals in stable bisexual groups where a few males monopolize reproduction over short periods. Increased testis size after evolutionary transitions to infanticide and secondary loss of infanticide in species with large testes suggest that female paternity dilution strategy efficiently reduces infanticide risk, and emphasizes the reversible nature of infanticide, which may appear and disappear according to the evolutionary arms race between the sexes. Although past studies failed to find comparative evidence for sexual counterstrategies to infanticide (7, 16), our sample spanning all mammals, combined with phylogenetically controlled analyses, and our focus on testis size to index female sexual promiscuity may explain divergent results.

Our analyses build on five decades of field observations to understand the distribution of species differences in the occurrence of infanticide by males across mammalian societies. The evolution of infanticide is largely determined by variation in the intensity of male-male contest competition and has not promoted major switches in mammalian social organization. However, it has affected the evolution of mating systems by increasing female sexual promiscuity as a paternity dilution strategy. These findings suggest that the distribution of infanticide across mammals is a consequence of contrasts in social and mating systems and closely reflects variation in the intensity of intra- and intersexual conflict.

#### REFERENCES AND NOTES

1. R. A. Palombit, in *The Evolution of Primate Societies*, J. C. Mitani, J. Call, P. Kappeler, R. A. Palombit, J. B. Silk, Eds. (Univ. of Chicago Press, Chicago, 2012), pp. 432–468.
2. S. B. Hrdy, *Folia Primatol. (Basel)* **22**, 19–58 (1974).
3. S. B. Hrdy, *Ethol. Sociobiol.* **1**, 13–40 (1979).



4. L. A. Ebersperger, *Biol. Rev. Camb. Philos. Soc.* **73**, 321–346 (1998).
5. C. Janson, C. P. Van Schaik, in *Infanticide by Males and Its Implications*, C. P. Van Schaik, C. H. Janson, Eds. (Cambridge Univ. Press, Cambridge, 2000), pp. 469–494.
6. C. L. Nunn, C. P. Van Schaik, in *Infanticide by Males and Its Implications*, C. P. Van Schaik, C. H. Janson, Eds. (Cambridge Univ. Press, Cambridge, 2000), pp. 388–412.
7. D. T. Blumstein, in *Infanticide by Males and Its Implications*, C. P. Van Schaik, C. H. Janson, Eds. (Cambridge Univ. Press, Cambridge, 2000), pp. 178–197.
8. C. P. van Schaik, in *Infanticide by Males and Its Implications*, C. P. van Schaik, C. H. Janson, Eds. (Cambridge Univ. Press, Cambridge, 2000), pp. 61–71.
9. D. Lukas, T. H. Clutton-Brock, *Science* **341**, 526–530 (2013).
10. C. Opie, Q. D. Atkinson, R. I. M. Dunbar, S. Shultz, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13328–13332 (2013).
11. G. R. Pradhan, C. Van Schaik, *Behaviour* **145**, 251–275 (2008).
12. J. A. Teichroeb, E. C. Wikberg, I. Badescu, L. J. Macdonald, P. Sicotte, *Behav. Ecol.* **23**, 1348–1359 (2012).
13. C. P. van Schaik, R. I. M. Dunbar, *Behaviour* **115**, 30–61 (1990).
14. C. Borries, T. Savini, A. Koenig, *Behav. Ecol. Sociobiol.* **65**, 685–693 (2011).
15. C. P. van Schaik, P. M. Kappeler, *Proc. Biol. Sci.* **264**, 1687–1694 (1997).
16. M. A. van Noordwijk, C. P. van Schaik, in *Infanticide by Males and Its Implications*, C. P. Van Schaik, C. H. Janson, Eds. (Cambridge Univ. Press, Cambridge, 2000), pp. 322–360.
17. J. O. Wolff, D. W. Macdonald, *Trends Ecol. Evol.* **19**, 127–134 (2004).
18. Materials and methods are available on Science Online.
19. S. A. Fritz, O. R. Bininda-Emonds, A. Purvis, *Ecol. Lett.* **12**, 538–549 (2009).
20. D. Lukas, T. Clutton-Brock, *Proc. Biol. Sci.* **281**, 20140418 (2014).
21. C. M. Crockett, C. H. Janson, in *Infanticide by Males and Its Implications*, C. P. Van Schaik, C. H. Janson, Eds. (Cambridge Univ. Press, Cambridge, 2000), pp. 75–98.
22. C. D. Soulsbury, *PLOS ONE* **5**, e9581 (2010).
23. A. H. Harcourt, P. H. Harvey, S. G. Larson, R. V. Short, *Nature* **293**, 55–57 (1981).
24. J. M. Beaulieu, D. C. Jhwueng, C. Boettiger, B. C. O'Meara, *Evolution* **66**, 2369–2383 (2012).

## ACKNOWLEDGMENTS

We are grateful to T. Clutton-Brock, P. Kappeler, G. Cowlishaw, D. Blumstein, and an anonymous reviewer for helpful comments on an earlier draft of this manuscript, and to the Large Animal Research Group for insightful discussions. D.L. received funding from the European Research Council (grant no. 294494-THCB2011) and E.H. from the Natural Environment Research Council (grant no. NE/RG53472). Data supporting this study are available in the supplementary materials.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/346/6211/841/suppl/DC1](http://www.sciencemag.org/content/346/6211/841/suppl/DC1)

Data and Methods

Supplementary Text

Tables S1 to S4

References (25–262)

10 June 2014; accepted 18 September 2014

10.1126/science.1257226

## LANDSCAPE HYDROLOGY

# The hydrological legacy of deforestation on global wetlands

C. Woodward,<sup>1\*</sup> J. Shulmeister,<sup>1</sup> J. Larsen,<sup>1</sup> G. E. Jacobsen,<sup>2</sup> A. Zawadzki<sup>2</sup>

Increased catchment erosion and nutrient loading are commonly recognized impacts of deforestation on global wetlands. In contrast, an increase in water availability in deforested catchments is well known in modern studies but is rarely considered when evaluating past human impacts. We used a Budyko water balance approach, a meta-analysis of global wetland response to deforestation, and paleoecological studies from Australasia to explore this issue. After complete deforestation, we demonstrated that water available to wetlands increases by up to 15% of annual precipitation. This can convert ephemeral swamps to permanent lakes or even create new wetlands. This effect is globally significant, with 9 to 12% of wetlands affected, including 20 to 40% of Ramsar wetlands, but is widely unrecognized because human impact studies rarely test for it.

The quality and quantity of water delivered to wetland ecosystems (including lakes) are key drivers of biodiversity and ecosystem function. There has been a major focus on the impact of catchment deforestation and changing land use on the quality of water draining to wetlands (e.g., nutrient and sediment loads) (1, 2), but there has been little attention paid to the legacy of deforestation on the water balance of wetlands. This is an important oversight, because hydrological studies have shown that cleared areas (grasslands and croplands) have about a one-third lower amount of actual evapotranspiration (ET) than do forests over a wide range of climates (3). This has the effect of increasing water yields from deforested catch-

ments (4–10) and should also increase the water available to wetlands (fig. S1). Deforestation may therefore alter wetland ecosystems by increasing water depth and persistence and may even create new wetlands. This mechanism has previously been attributed to the onset of peat formation in Western Europe ~5000 years ago (11), when humans first began large-scale deforestation in these catchments. How this impact should be manifest in terms of wetland hydrology and ecology in the wider spectrum of global deforestation and climate regimes has so far escaped attention.

We first explored this potential impact conceptually using the Budyko hypothesis (12), which expresses the partitioning of water within catchments according to the annual supply and demand of atmospheric water. This widely used approach normalizes atmospheric demand [potential evaporation (EP)] and actual evaporation (E) to supply (precipitation, P) and hence is more useful than methods that only estimate absolute changes in EP (13). It assumes that over annual time scales,

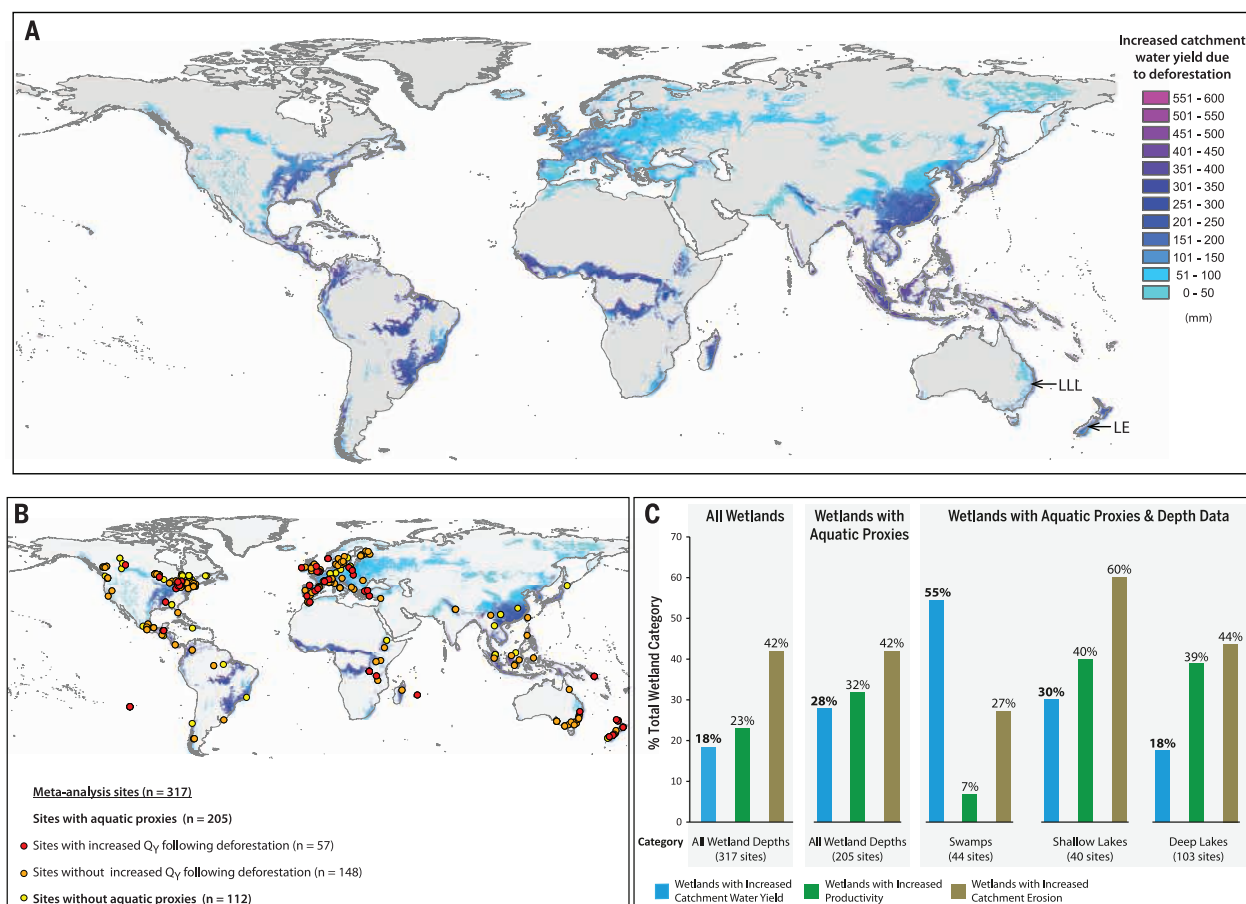
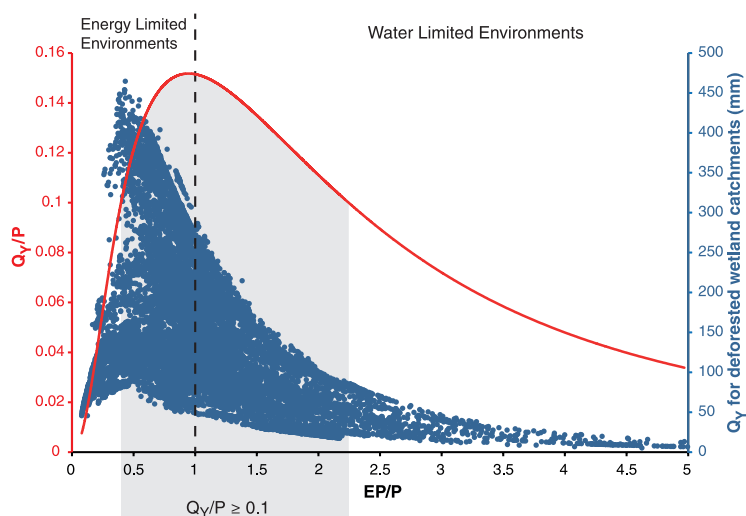
changes in soil and groundwater storages are small and can be related to catchment water yield ( $Q_Y$ ) and other effects such as deforestation (3). We used a well-defined relationship for the reduction in ET due to complete catchment deforestation (forest to grassland or cropland conversion) in the context of the Budyko hypothesis to predict normalized increases in water yield ( $Q_Y/P$ ) (14). We find that  $Q_Y/P$  should increase slowly in water-limited environments ( $EP/P > 1$ ), reach a maximum close to the boundary between water- and energy-limited environments ( $EP/P \approx 1$ ), and decline very sharply within energy-limited environments ( $EP/P < 1$ ) (Fig. 1). Extra water available to wetlands within deforested catchments is equivalent to 10 to 15% of annual rainfall when catchment  $EP/P$  is within the wide range of 0.4 to 2.2.

We used the relationship between water yield and  $EP/P$  derived from the Budyko water balance analysis with a global wetland database of ~245,000 wetlands to investigate the effect of deforestation on global wetland hydrology (14). There is considerable uncertainty surrounding the distribution of forest in the absence of clearance by humans (15), so we used a range of potential natural forest cover estimates in our calculations (fig. S4 and table S1). Nine to 12% of global wetlands, including 20 to 40% of Ramsar wetlands, are situated in formerly forested areas. Water yield has increased by up to 460 mm in these catchments (Fig. 1, figs. S5 to S7, and table S2). In catchments with extremely high P, local water gain can exceed 600 mm (Fig. 2A). How this translates to hydrological changes in individual wetlands depends on the ratio of the catchment to wetland area and hypsometry. Wetlands with higher catchment-to-lake area ratios will have larger impacts, as will shallow and/or ephemeral systems, where these changes are amplified (16).

To expand on the Budyko analysis, we performed a global meta-analysis on 317 wetland sites where Holocene deforestation and changes in the wetlands have been recorded (14) (Fig. 2B and database S1). A significant increase in

<sup>1</sup>School of Geography, Planning and Environmental Management, University of Queensland, Chamberlain Building, St Lucia Campus, Brisbane, 4072 Queensland, Australia. <sup>2</sup>Institute for Environmental Research, Australian Nuclear Science and Technology Organisation, Locked Bag 2001, Kirrawee DC 2232, New South Wales, Australia.  
\*Corresponding author. E-mail: c.woodward1@uq.edu.au

**Fig. 1. Predicted changes in catchment hydrology due to deforestation based on the Budyko water balance analysis (14).** The potential increase in catchment water yield after deforestation ( $Q_Y$ ) is calculated as a fraction of the catchment  $P$  and is depicted as  $Q_Y/P$  (red curve). Values for  $Q_Y/P$  depend on the ratio of  $EP$  to  $P$  in the catchment.  $Q_Y/P$  is largest ( $\geq 0.1$ ) in the gray shaded area, where  $EP/P$  is between 0.4 and 2.2. The relationship between  $EP/P$  and  $Q_Y/P$  was used to calculate the actual increase in  $Q_Y$  for wetlands in deforested areas depicted in Fig. 2A (blue circles,  $n = 25,264$  wetlands; see fig. S6 for the wetland locations).  $Q_Y/P$  translates into  $\leq 460$  mm of extra water yield in these catchments (mean = 125 mm, SD = 85). The peak in calculated catchment water yield values lies to the left of the  $Q_Y/P$  peak, because actual  $P$  values are much higher in energy-limited environments (fig. S5). Where the initial forest cover was naturally low or deforestation is not complete,  $Q_Y$  may be an overestimate.



**Fig. 2. Global map of areas susceptible to hydrological change after deforestation and a meta-analysis of Holocene records of catchment deforestation and wetland change.** (A) Potential increase in  $Q_Y$  in deforested areas. This is derived from a possible area of forest cover without anthropogenic deforestation and the results of the Budyko water balance analysis depicted in Fig. 1 (14). In extreme circumstances, this increase in water yield may be as large as 600 mm. The locations of the paleoecological records presented in this paper—Little Llangothlin Lagoon (LLL) and Lake Emma (LE)—are shown. (B) Colored circles indicate the localities of 317 studies in

the meta-analysis of Holocene wetland response to catchment deforestation (database S1). Sites are color-coded based on whether the studies used aquatic proxies (such as diatoms) and whether a hydrological response to deforestation was detected. Hydrological responses to deforestation include wetland formation and increased water depth and/or persistence. (C) A summary of the response of wetlands in the meta-analysis to deforestation. The detection of post-deforestation increases in water yield was higher in the subset of wetlands that used aquatic proxies. Hydrological changes are more important than eutrophication and increased sedimentation rates in swamps.



catchment water yield and change in wetland hydrology was recorded in 18% (57) of these sites (Fig. 2C). This understates the significance of this phenomenon, because 112 sites in the database did not include analyses for proxies that might enable hydrological changes to be detected (such as aquatic pollen or diatoms) (Fig. 2B). When only sites with suitable aquatic proxies are considered, the number of clearly

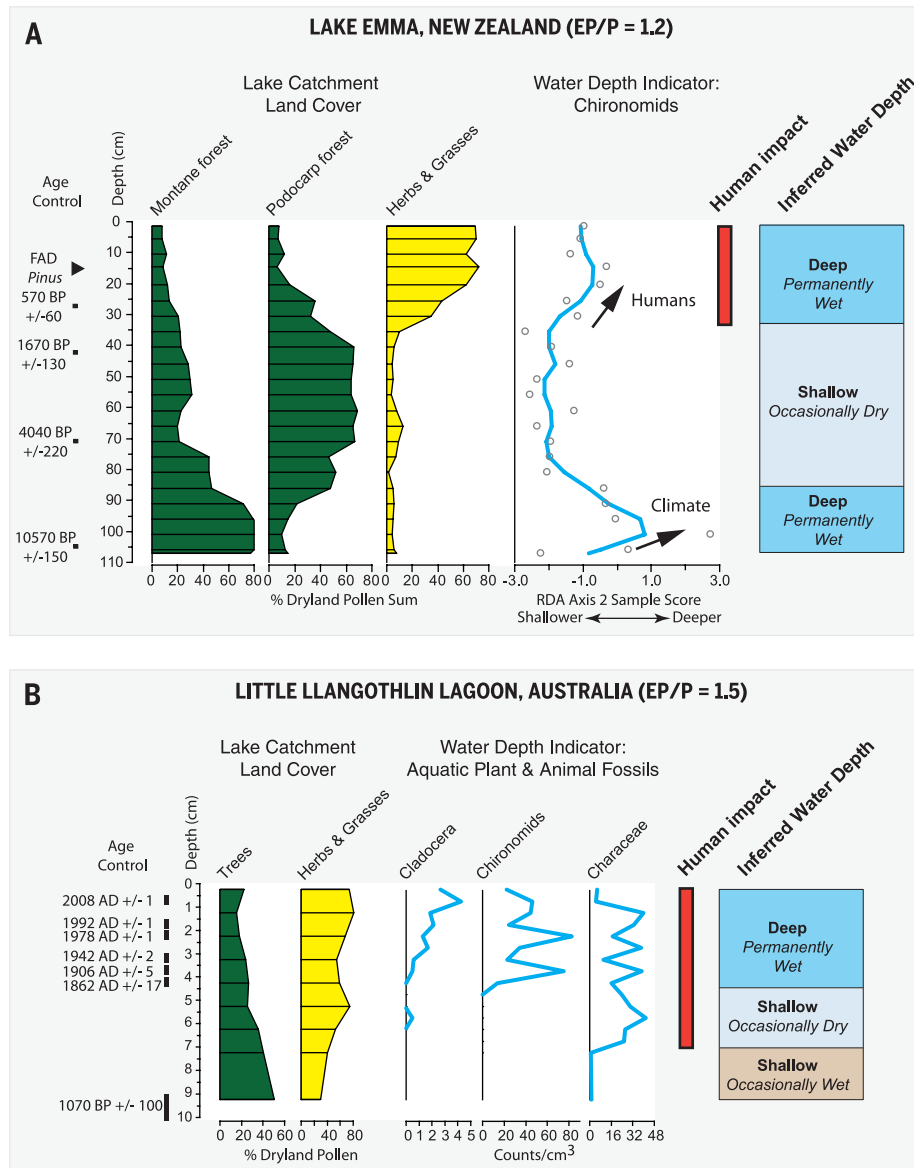
affected wetlands increases to 28%. This approaches the percentage of wetlands where better-recognized post-deforestation changes, such as eutrophication or increases in sedimentation rate, have been reported (Fig. 2C). We suspect that this is still an underestimation of the actual magnitude of the hydrological legacy of deforestation, because most of the studies in the database were not designed to examine changes in hydrology.

Obvious changes (such as wetland formation or the conversion of swamps to lakes) are more likely to be detected after the fact, whereas more subtle changes (such as increases in depth in deep lakes) are unlikely to be observed. This argument is supported by the fact that detection of a hydrological change was more common in swamps than lakes (Fig. 2C).

We provide two case studies from Australasia to illustrate how post-deforestation hydrological changes can result in a major alteration of wetland ecology. These studies also demonstrate how this effect can be missed by paleoecological studies, even if they do use suitable proxies to detect hydrological change. We examined a paleoecological record from New Zealand and one from Australia in areas where eutrophication and increased sedimentation rates are assumed to be the major effects of deforestation (17, 18). We reconstructed hydrological changes (14) from a wetland on the eastern side of the South Island, New Zealand (Lake Emma), and on the New England Tablelands of New South Wales in Australia (Little Llangothlin Lagoon) (Fig. 2A, fig. S8, and fig. S11). Changes in fossil pollen (Fig. 3A) indicate widespread deforestation near Lake Emma after initial human (Māori) settlement in ~1250 CE (19). At Little Llangothlin Lagoon (Fig. 3B), European deforestation started in the 1860s and peaked in the 1870s (20). Both sites have remained deforested since then.

We used nonbiting midge (Chironomidae) fossils to reconstruct relative changes in water depth during the late Holocene in Lake Emma (Fig. 3A) (14). The water level increased after Māori deforestation ~700 years ago to levels not experienced in this lake for many thousands of years (Fig. 3A). In addition, ephemeral lake beds at Spider Lakes close to this study site (fig. S8) have a basal age of  $480 \pm 180$  calendar years before the present, indicating that these now shallow (<1 m deep) intermittent lakes were non-swamp hollows before Māori deforestation. In Little Llangothlin Lagoon, high concentrations of aquatic plant and animal remains do not appear in the record until after European impact (Fig. 3B). In particular, aquatic invertebrate fossils (cladocera and chironomids) are not consistently present until after ~1860 CE (Fig. 3B), indicating permanent water in Little Llangothlin Lagoon for the first time in several thousand years. Records from elsewhere in Little Llangothlin Lagoon also support the reconstruction from this site (fig. S12). In summary, our records demonstrate that the main impact of deforestation at these sites is an increase in available water to the wetlands and not increased catchment erosion or nutrient loading.

Available palaeoclimate records from Australia and New Zealand show no evidence of wetter climates, either after Māori settlement of New Zealand or European settlement of eastern Australia (21–25). In addition, the Budyko approach highlights the fact that small changes in climate (P and EP) are only significant for notably water- or energy- limited environments (i.e., very high or low EP/P), where wetlands are more sensitive



**Fig. 3. Paleoecological records of human settlement, land use, and hydrological change from New Zealand and Australia. (A)** Lake Emma, New Zealand. The pollen record indicates the presence of forest in the lake catchment up until Māori deforestation ~700 years ago (19). Deforestation reduced ET in the lake catchment, increased water yield, and therefore increased lake water depth. A deep-water period earlier in the Holocene (~10,000 to 6000 years ago) was driven by climate. FAD, first appearance datum; BP, before the present. **(B)** Little Llangothlin Lagoon, Australia. The lagoon was mostly dry for at least 1000 years before European deforestation (~1860 to 1870 CE). The appearance of Characeae (charophytes) after 1860 indicates the increased persistence of water in the lagoon. The presence of aquatic invertebrate remains (cladocera and chironomids) marks the onset of permanent standing water. Longer records from elsewhere in the lagoon confirm the absence of standing water in the lagoon before deforestation for many thousands of years (fig. S12).

to changes in P and EP, respectively. In practice, this means that a major change in climate would be required to replicate the impacts of deforestation within the more moderate EP/P ranges (EP/P ~ 0.4 to 2.2) in these wetland catchments. Therefore, the observed increases in water yield in the Australasian paleoecological records are better explained by deforestation rather than climate change.

We conclude that increased water yield is a critical factor in the evolution and maintenance of wetlands in deforested regions. These findings have implications for the way we study and manage global lakes and wetlands. Studies should be designed to examine baseline hydrological conditions in deforested wetlands and not just changes in sedimentation rate or trophic status. Reconstructions of past wetland environments have largely failed to recognize enhanced water availability as a consequence of human disturbance. The additional water yield may result in much more water being available for ecosystem functioning, particularly in subhumid regions. From a management point of view, the reforestation of wetland catchments may drastically alter this water balance, and in some cases, the protected wetland will cease to exist. Therefore, wetland management practices and policies that include catchment reforestation as a stated goal also need to consider the potential hydrological consequences before being implemented.

#### REFERENCES AND NOTES

- M. Søndergaard, E. Jeppesen, *J. Appl. Ecol.* **44**, 1089–1094 (2007).
- M. M. Brinson, A. I. Malvarez, *Environ. Conserv.* **29**, 115–133 (2002).
- L. Zhang, W. R. Dawes, G. R. Walker, *Water Resour. Res.* **37**, 701–708 (2001).
- J. M. Bosch, J. D. Hewlett, *J. Hydrol. (Amst.)* **55**, 3–23 (1982).
- A. E. Brown, L. Zhang, T. A. McMahon, A. W. Western, R. A. Vertessy, *J. Hydrol. (Amst.)* **310**, 28–61 (2005).
- L. A. Bruijnzeel, *J. Trop. For. Sci.* **1**, 229–243 (1988).
- K. A. Farley, E. G. Jobbágy, R. B. Jackson, *Glob. Change Biol.* **11**, 1565–1576 (2005).
- J. W. Hornbeck, M. B. Adams, E. S. Corbett, E. S. Verry, J. A. Lynch, *J. Hydrol. (Amst.)* **150**, 323–344 (1993).
- V. Sahin, M. J. Hall, *J. Hydrol. (Amst.)* **178**, 293–309 (1996).
- R. Silberstein, A. Adhitya, C. Dabrowski, *Changes in Flood Flows, Saturated Area and Salinity Associated with Forest Clearing for Agriculture* (Technical Report 03/01, Cooperative Research Centre for Catchment Hydrology, Monash University, Victoria, Australia, 2003).
- H. Cubizolle et al., *Quat. Int.* **251**, 77–96 (2012).
- M. I. Budyko, *Climate and Life* (Academic Press, New York, 1974).
- S. M. Sterling, A. Ducharme, J. Polcher, *Nat. Clim. Change* **3**, 385–390 (2012).
- Materials and methods are available as supplementary materials on Science Online.
- W. J. Bond, F. I. Woodward, G. F. Midgley, *New Phytol.* **165**, 525–538 (2005).
- K. Szesztay, *Hydrol. Sci. Bull.* **19**, 73–84 (2010).
- D. B. McWethy et al., *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21343–21348 (2010).
- S. J. Gale, R. J. Haworth, P. C. Pisanu, *Quat. Sci. Rev.* **14**, 395–408 (1995).
- C. A. Woodward, J. Shulmeister, A. Zawadzki, G. Jacobsen, *Holocene* **24**, 668–678 (2014).
- B. J. Stubbs, *Environ. Hist.* **4**, 145–165 (1998).
- P. W. Williams, D. N. T. King, J.-X. Zhao, K. D. Collerson, *Earth Planet. Sci. Lett.* **230**, 301–317 (2005).
- M. G. Macklin, I. C. Fuller, A. F. Jones, M. Bebbington, *Geology* **40**, 775–778 (2012).
- M. J. Page et al., *Mar. Geol.* **270**, 30–44 (2010).
- J. Gergis et al., *Clim. Change* **111**, 923–944 (2012).
- T. R. Vance, T. D. van Ommen, M. A. J. Curran, C. T. Plummer, A. D. Moy, *J. Clim.* **26**, 710–725 (2013).

#### ACKNOWLEDGMENTS

This research was part of an Australian Research Council Discovery Project (DP 110103081) and was also partly supported by the Australian Institute for Nuclear Science and Engineering (grants ALNGRA11068 and ALNGRA12071). The database used for the meta-analysis of Holocene records of catchment deforestation

and wetland change can be downloaded from the Dryad Digital Repository at <http://dx.doi.org/10.5061/dryad.m62gj>.

#### SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/346/6211/844/suppl/DC1](http://www.sciencemag.org/content/346/6211/844/suppl/DC1)  
Supplementary Text  
Figs. S1 to S13  
Tables S1 and S2  
References (26–43)  
Databases S1 to S3

27 August 2014; accepted 17 October 2014  
10.1126/science.1260510

#### PALEOCEANOGRAPHY

## Antarctic role in Northern Hemisphere glaciation

Stella C. Woodard,<sup>1\*</sup> Yair Rosenthal,<sup>1,2</sup> Kenneth G. Miller,<sup>2</sup> James D. Wright,<sup>2</sup> Beverly K. Chiu,<sup>2</sup> Kira T. Lawrence<sup>3</sup>

Earth's climate underwent a major transition from the warmth of the late Pliocene, when global surface temperatures were ~2° to 3°C higher than today, to extensive Northern Hemisphere glaciation (NHG) ~2.73 million years ago (Ma). We show that North Pacific deep waters were substantially colder (4°C) and probably fresher than the North Atlantic Deep Water before the intensification of NHG. At ~2.73 Ma, the Atlantic-Pacific temperature gradient was reduced to <1°C, suggesting the initiation of stronger heat transfer from the North Atlantic to the deep Pacific. We posit that increased glaciation of Antarctica, deduced from the 21 ± 10-meter sea-level fall from 3.15 to 2.75 Ma, and the development of a strong polar halocline fundamentally altered deep ocean circulation, which enhanced interhemispheric heat and salt transport, thereby contributing to NHG.

**B**enthic foraminiferal oxygen isotope ( $\delta^{18}\text{O}_{\text{bf}}$ ) records trend toward higher values throughout the late Pliocene warm period (1), culminating in higher amplitude of glacial-interglacial (G-IG)  $\delta^{18}\text{O}$  variability (Fig. 1) associated with a dramatic increase of ice-rafted debris in sediments around the Arctic ~2.73 million years ago (Ma) [marine isotope stage (MIS) G6], and signaling a major intensification of Northern Hemisphere glaciation (NHG) (2–4). During the same period, proxy records indicate a long-term sea surface temperature (SST) cooling in most regions (5–8), which is mimicked in bottom water temperature (BWT) records from the deep North Atlantic (9–11). Here, we reconstruct the BWT history of the deep Pacific at 2- to 5-thousand year (ky) resolution from Ocean Drilling Program (ODP) site 1208 (36.13°N, 158.20°W, 3350 m water depth) (fig. S1) on Shatsky Rise in the northwest Pacific Ocean (12), and we determine ice volume changes during the 400-ky interval preceding the NHG intensification (3.15 to 2.75 Ma, henceforth “pre-NHG”).

The Pliocene  $\delta^{18}\text{O}_{\text{bf}}$  record at site 1208 is consistent with the global  $\delta^{18}\text{O}_{\text{bf}}$  stack (LR04) (3) (Fig. 1). Beginning at glacial MIS G6, the amplitude of G-IG  $\delta^{18}\text{O}_{\text{bf}}$  variability increases from ~0.5 to >0.8 per mil (‰) with more extreme glacial maxima and interglacial minima (fig. S2).

The BWT record at site 1208 was reconstructed from Mg/Ca measurements of the infaunal foraminifera, *Uvigerina* spp. We determined a Mg/Ca temperature sensitivity ( $0.067 \pm 0.016 \text{ mmol mol}^{-1} \text{ } ^\circ\text{C}^{-1}$ ) for this species by calibrating the Holocene–Last Glacial Maximum (LGM)  $\Delta(\text{Mg/Ca})$  to the  $\Delta\delta^{18}\text{O}_{\text{bf}}$  at site 1208 (12, 13). Note that unlike the epifaunal species *Cibicides wuellerstorfi*, which was used to generate a North Atlantic BWT record (9), *Uvigerina* is apparently not affected by seawater carbonate ion saturation, which can alter Mg/Ca independently of temperature (12, 14). Pacific BWTs exhibit an average Pliocene G-IG range of  $\sim 2.1 \pm 0.8^\circ\text{C}$  (1 SD), similar to the average Holocene–LGM range of  $2.6^\circ\text{C}$  (12) and in contrast with oscillations of  $\delta^{18}\text{O}_{\text{bf}}$  (Fig. 1), suggesting that the orbital-scale temperature variability of the Pliocene Pacific was somewhat decoupled from the ice volume. A notable change occurred ~2.73 Ma, when, following ~4°C deglacial warming, average Pacific BWT became ~1.5°C warmer and G-IG temperature changes exhibited slightly lower amplitude (Fig. 1 and fig. S3).

Spectral analysis shows that pre-NHG our  $\delta^{18}\text{O}_{\text{bf}}$  record is dominated by 41-ky pacing, consistent

<sup>1</sup>Department of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, NJ 08901, USA.

<sup>2</sup>Department of Earth and Planetary Sciences, Rutgers University, 610 Taylor Road, Piscataway, NJ 08845, USA.

<sup>3</sup>Department of Geology and Environmental Geosciences, Lafayette College, 730 High Street, Easton, PA 18042, USA.

\*Corresponding author. E-mail: [woodard@marine.rutgers.edu](mailto:woodard@marine.rutgers.edu)



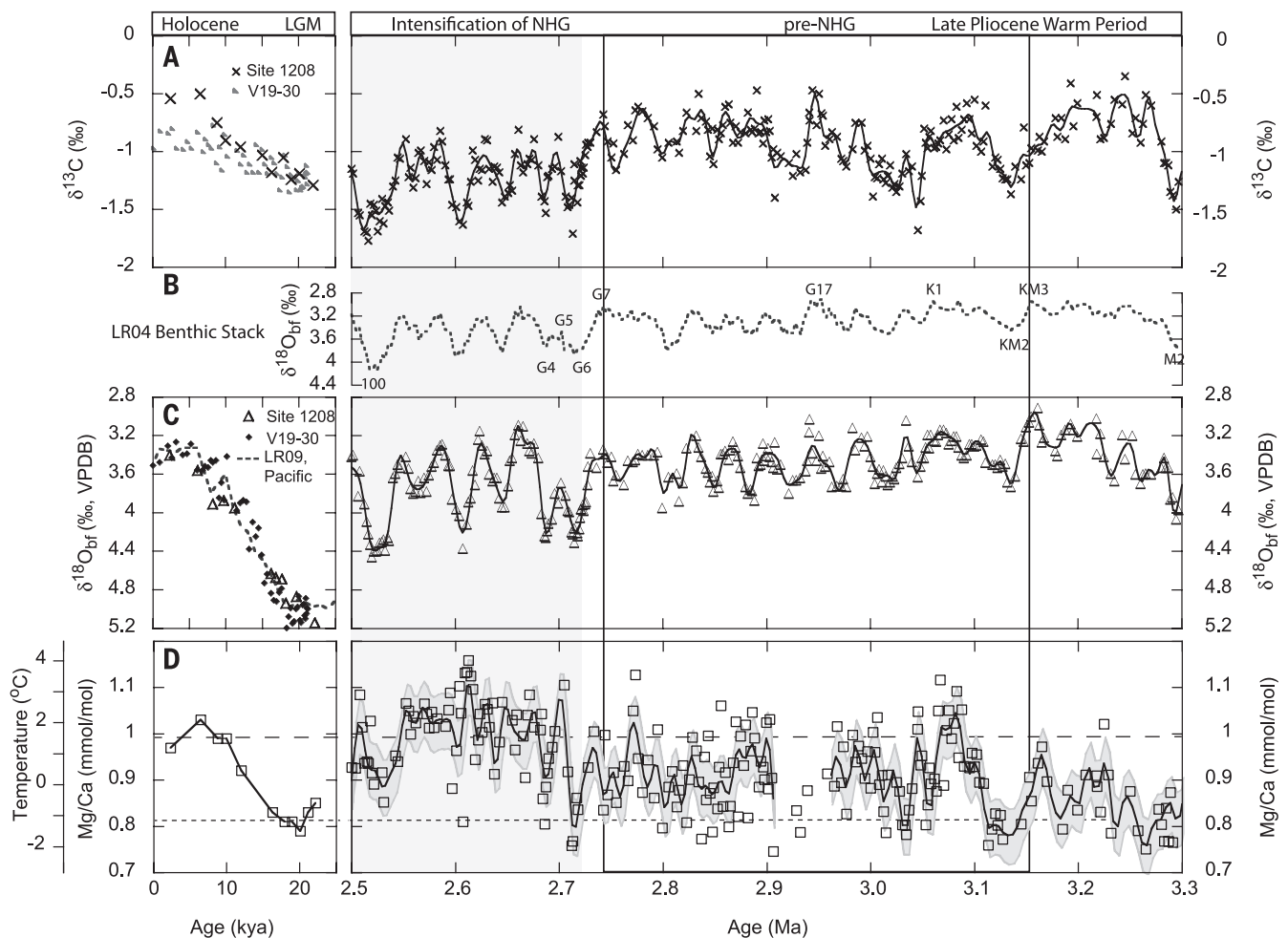
with a strong response to obliquity variations, which is attributed to ice sheet dynamics (3). In contrast, Pacific BWT and benthic  $\delta^{13}\text{C}$  exhibit strong 100-ky periodicity during that interval (fig. S4), suggesting that BWT and circulation responded differently to orbital forcing than ice. The appearance of a secondary 100-ky beat in the latter part of our  $\delta^{18}\text{O}_{\text{bf}}$  record provides evidence for stronger synchronization in the response of these climate-system components to orbital forcing as NHG intensified (12).

Site 1208 BWTs show no significant long-term trend over the 400-ky pre-NHG interval (Fig. 2). Therefore, we interpret the significant  $\delta^{18}\text{O}_{\text{bf}}$  increase [ $0.21 \pm 0.04\text{‰}$  (1 SE)] at site 1208 over this interval (Fig. 2) to primarily reflect an increase in the extent of continental ice (12). Assuming an average  $0.1 \pm 0.02\text{‰}$   $\Delta\delta^{18}\text{O}$  per 10 m of sea level (15), we estimate a  $21 \pm 10\text{-m}$  (2 SE) sea-level equivalent of “permanent” ice growth (12) pre-NHG. Because seawater Mg/Ca was probably

$\sim 20 \pm 10\%$  lower than modern levels during the Pliocene [e.g., (16)], BWTs at site 1208 might have been higher than present temperatures by  $\sim 1^\circ\text{C}$  (12) (fig. S3). However, this uncertainty has a negligible effect on our estimates of sea-level change and the G-IG variability of BWT and  $\delta^{18}\text{O}_{\text{sw}}$  during the late Pliocene interval leading to the NHG intensification (12) (fig. S5). Using the long-term trends also minimizes uncertainties due to inaccurate temperature corrections or differences in the primary orbital beats of the isotope and temperature records. Thus, our reconstruction of continental ice provides only information regarding mean sea-level position and does not reflect glacial lowstands or interglacial highstands, which presumably occurred at orbital frequencies over this 400-ky interval (12).

We compare the Pacific with North Atlantic Deep Sea Drilling Project (DSDP) site 607 ( $41^\circ\text{N}$ ,  $32^\circ\text{W}$ , 3427-m water depth) records (9) and find that pre-NHG Atlantic BWTs generated using

*C. wuellerstorfi* are, on average,  $4 \pm 1.4^\circ\text{C}$  warmer than at site 1208 (9) (Fig. 2). New discrete measurements of *Uvigerina* Mg/Ca from site 607 range from 1.1 to 1.6  $\text{mmol mol}^{-1}$ , well above values recorded by *Uvigerina* in the Pacific at the same time (fig. S6), supporting the interpretation of warm North Atlantic BWTs during the Pliocene (12). Over the pre-NHG interval, site 607  $\delta^{18}\text{O}_{\text{bf}}$  records a  $0.48 \pm 0.04\text{‰}$  increase, more than double that observed at site 1208, and BWTs cool by  $\sim 1.3 \pm 0.4^\circ\text{C}$  (Fig. 2). After accounting for BWT cooling, we estimate a  $\sim 15 \pm 14\text{-m}$  (1 SE) sea-level equivalent of “permanent” ice growth from the residual  $\Delta\delta^{18}\text{O}_{\text{bf}}$  at site 607 (12), consistent with the estimate from site 1208. As the uncertainty in the Pacific sea-level record is less than the Atlantic estimate due to lack of temperature correction, we conclude that baseline sea-level fall very likely (>95% confidence) exceeded 11 m and possibly reached 31 m pre-NHG.



**Fig. 1. Stable isotope and Mg/Ca data for the deep Pacific Ocean, site 1208.** Records span the Holocene-LGM [0 to 25 thousand years ago (kya in figure), left] and the late Pliocene (3.3 to 2.5 Ma, right). Holocene-LGM data from V19-30 (eastern tropical Pacific) (38) and the Pacific  $\delta^{18}\text{O}$  benthic stack (39) are included for comparison. (A) *Uvigerina* spp.  $\delta^{13}\text{C}$  records from the Pacific Ocean. (B) Global benthic  $\delta^{18}\text{O}$  stack (LR04) (3). (C) *Uvigerina* spp.  $\delta^{18}\text{O}_{\text{bf}}$  record from site 1208. (D) *Uvigerina* spp. Mg/Ca record from site

1208. The temperature scale is provided for reference. Raw data (points) were smoothed using a 15-ky low-pass filter (solid lines); gray shading indicates 1 SE of temperature estimate. The rectangle vertically spanning (A) to (D) identifies the pre-NHG interval, and the gray vertically shaded area highlights the period after NHG intensification. Important Pliocene MISs are labeled on the LR04 benthic  $\delta^{18}\text{O}$  stack. VPDB, Vienna Pee Dee belemnite.

The gradual increase in Pliocene  $\delta^{18}\text{O}_{\text{bf}}$  records has previously been attributed to a prolonged onset of NHG (17, 18). Ice-rafted debris provides evidence of ice on Greenland since the Miocene and an augmented ice sheet extent  $\sim 3.3$  Ma (19). Although tectonic uplift in high northern latitudes surrounding the Barents Sea likely promoted regional ice growth as early as 4 Ma (20), the alpine glaciers were not large enough to reach sea level until MIS G6 (20) and so probably account for negligible sea-level equivalent ice pre-NHG (12). In light of the little evidence for widespread glaciers in the Northern Hemisphere until  $\sim 2.73$  Ma (4, 19, 20), we suggest that at least half of the  $21 \pm 10$ -m sea-level fall pre-NHG was due to ice added to Antarctica. Furthermore, because the Greenland Ice Sheet and West Antarctic Ice Sheet can accommodate at most a  $\sim 12$ -m ice equivalent of sea level (21, 22), it is likely that the extent of the East Antarctic Ice Sheet was less than modern levels during the late Pliocene, consistent with records of a dynamic Antarctic margin (23–26).

The expansion of Antarctic ice sheets during the late Pliocene appears to surpass an important climatological threshold, which contributed to the intensification of NHG. The present deep Pacific  $\delta^{18}\text{O}_{\text{sw}}$  value,  $\sim -0.15\text{‰}$ , reflects the mixture of North Atlantic Deep Water (NADW) ( $\delta^{18}\text{O}_{\text{sw}} \sim -0.2\text{‰}$ ) and Antarctic Bottom Water (AABW) ( $\delta^{18}\text{O}_{\text{sw}} \sim -0.5$  to  $-0.3\text{‰}$ ) forming Circumpolar Deep Water (CDW) (27). The pre-NHG deep

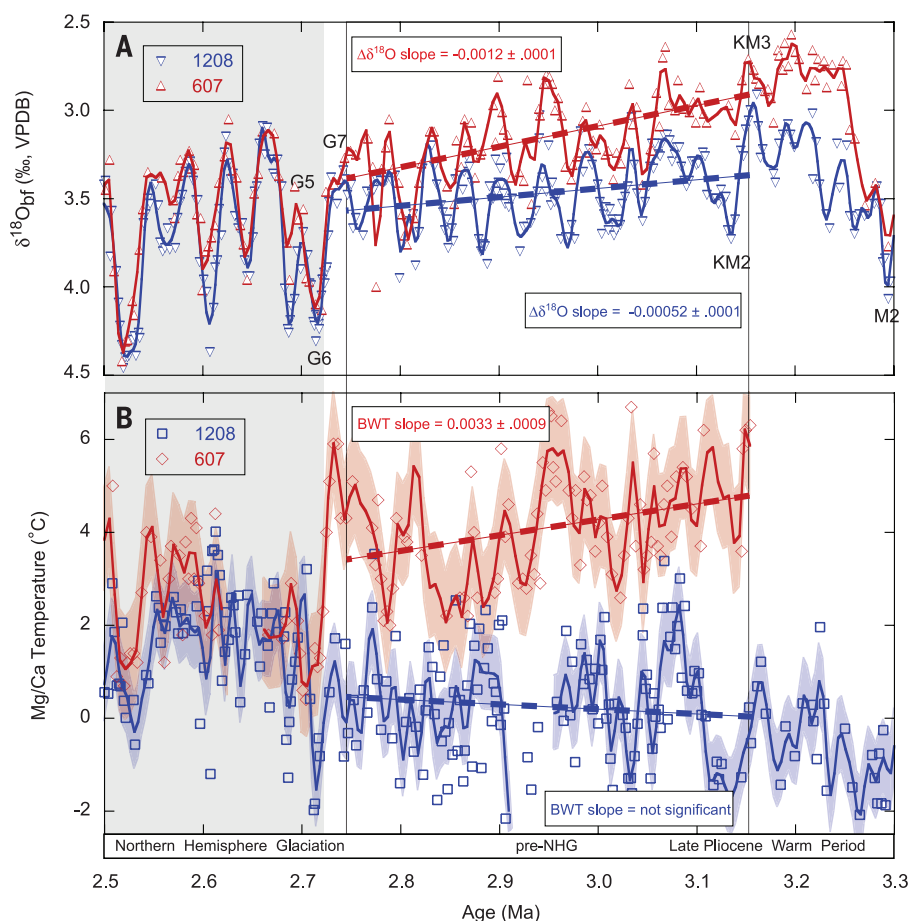
Pacific  $\delta^{18}\text{O}_{\text{sw}}$  was similar to modern AABW, averaging  $-0.43\text{‰}$  and fluctuating by as much as 0.4 to 0.8‰ on G-IG time scales (Fig. 3). During this interval, Pacific  $\delta^{18}\text{O}_{\text{sw}}$  reached near-modern values only during glacial periods, with much lower values during interglacials. The  $\delta^{18}\text{O}_{\text{sw}}$  reconstruction at site 607 shows no significant long-term change over the entire interval for which there is data overlapping with our Pacific record (9) (Fig. 3). Although the site 607 Pliocene G-IG amplitude of  $\Delta\delta^{18}\text{O}_{\text{sw}}$  was similar to that of the Pacific, average  $\delta^{18}\text{O}_{\text{sw}}$  was higher (0.28‰), with interglacial values not much different from those of the modern North Atlantic. Thus, the low Pacific  $\delta^{18}\text{O}_{\text{sw}}$  value cannot be attributed to lower continental ice volume but must reflect a hydrographic difference.

The persistence of an average  $\sim 0.7\text{‰}$   $\delta^{18}\text{O}_{\text{sw}}$  gradient between sites 607 and 1208 (Fig. 3B) suggests that the two sites were not ventilated by the same water mass pre-NHG. During the same interval, Pacific  $\delta^{18}\text{O}_{\text{bf}}$  and BWTs were offset from the North Atlantic (Fig. 2). The relatively faster long-term  $\delta^{18}\text{O}_{\text{bf}}$  increase at site 607 compared with site 1208 brings the records together by MIS G6 (Fig. 2A), and the Atlantic-Pacific BWT gradient was reduced from  $\sim 4^\circ\text{C}$  to  $<1^\circ\text{C}$  over one IG-G cycle [MIS G6 to MIS G5 (Fig. 2B)] by cooling at site 607 and subsequent warming at site 1208. We interpret these data to reflect hydrographic changes, which initiated stronger heat

and salt transfer from the North Atlantic to the Pacific due to a change in deep ocean circulation, perhaps setting up a CDW mixing pattern similar to that of the modern day.

Today, most AABW production occurs in the Weddell Sea (28), increasing opportunities for mixing between northern- and southern-sourced deep water masses due to its location in the Atlantic sector of the Southern Ocean. If the formation regions for Pliocene southern-sourced deep water were altered, an entirely different pattern of mixing and CDW formation was possible. Detailed reconstructions of the late Pliocene from the proximal ANDRILL AND-1B core indicate numerous episodes of near-total West Antarctic Ice Sheet collapse (25), and diatom assemblages suggest that open marine conditions were accompanied by deep mixing in the Ross Sea (26). Increased Pliocene deep water production in the Pacific-Indian sector of the Southern Ocean (e.g., Prydz Bay, Ross Sea, Amundsen Sea) potentially allowed AABW to influence the deep Pacific more directly, consistent with recent model results (29).

Past studies, which attribute Pliocene warmth to increased oceanic heat transport via the thermohaline conveyor (30–32), cite weak meridional  $\delta^{13}\text{C}$  gradients (30, 31) and increased North Atlantic sediment drift accumulation (32) as evidence of enhanced NADW production. However, their interpretation of ocean circulation based on  $\delta^{13}\text{C}$  hinges upon the use of modern NADW



**Fig. 2. Oxygen stable isotope and BWT records for the late Pliocene.** Benthic  $\delta^{18}\text{O}_{\text{bf}}$  (A) and BWTs (B) comparing Pacific (blue, this study) with the North Atlantic [red, (9)] records. The gray shaded area highlights the interval after Pacific and Atlantic  $\delta^{18}\text{O}_{\text{bf}}$  and BWT records come together around MIS G6. Long-term trends spanning the 400-ky pre-NHG interval (3.15 to 2.75 Ma) are indicated. Raw data (points) are smoothed by 15-ky low-pass filter (solid lines). Blue (site 1208) and red (site 607) shading gives 1 SE of filtered BWTs. Dashed lines are least-squares regressions. The slopes of significant trends are shown in boxes. MISs discussed in the text are labeled in (A).

and AABW end-members that are related to surface nutrient content and atmospheric exchange and thus may not reflect Pliocene values. Evidence for poorly stratified (33, 34) and more productive (24, 25) Southern Ocean surface waters with less sea ice during the Pliocene makes a higher preformed  $\delta^{13}\text{C}$  AABW end-member plausible (35). Although there is evidence for NADW production during the late Pliocene based on drift deposits (32), we have no direct way to quantify deep water formation relative to modern rates or to AABW. Both AABW and NADW production may have increased simultaneously, as suggested by modeling (29), with the end result of a lower contribution of NADW to CDW.

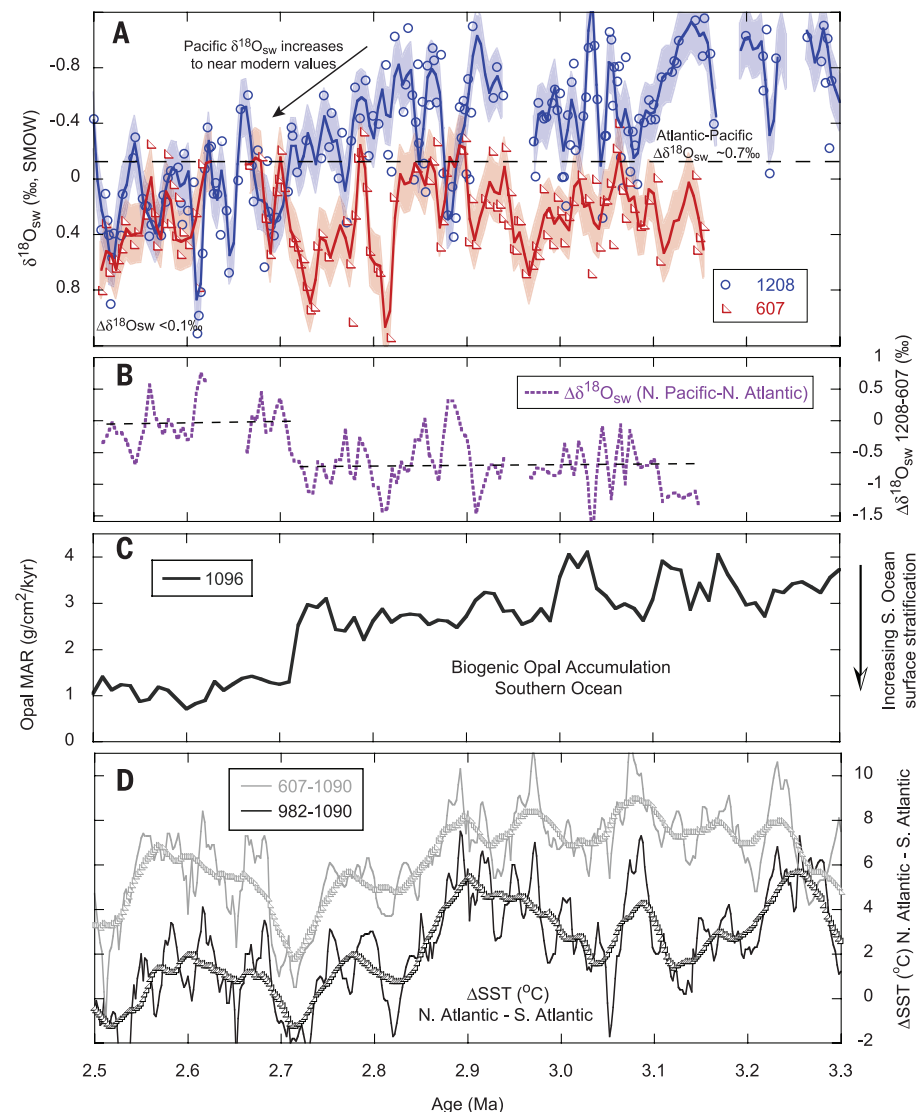
Regardless, the absence of a long-term temperature trend in our Pacific BWT record over the period when NADW cooled by  $\sim 1.3^\circ\text{C}$  attests to inefficient heat transport from the North Atlantic to the deep Pacific via the Southern Ocean before 2.73 Ma. We contend that NADW was not a major contributor to CDW and that deep water formed in the Southern Ocean was the primary source of deep water to the Pacific pre-NHG. This interpretation is also supported by a  $\sim 0.6\text{‰}$  increase in site 1208  $\delta^{18}\text{O}_{\text{sw}}$ , which effectively eliminates the Pacific-Atlantic  $\delta^{18}\text{O}_{\text{sw}}$  gradient by  $\sim 2.7$  Ma (Fig. 3). The proposed change in deep ocean circulation was almost certainly tied to high-latitude processes that altered the temper-

ature or salinity of either the northern or southern deep water mass, or both. Presumably, high-latitude cooling steepened equator-to-pole temperature gradients throughout the Pliocene (6, 7) and triggered the expansion of Antarctic glaciation in the Southern Hemisphere (26), as supported by our sea-level estimates from site 1208. Dust records suggest a northward shift in Southern Hemisphere wind belts about the time of NHG intensification (36), which would have allowed expansion of sea ice and promoted the development of a stronger halocline (34), reducing deep water ventilation around Antarctica (26, 35).

Stronger and more permanent stratification of Southern Ocean surface waters is indicated by decreased opal accumulation beginning around 2.73 Ma (33) (Fig. 3C). Such stratification probably reduced heat exchange between the atmosphere and upwelling Southern Ocean waters [e.g., (37)], trapping heat transported via NADW in the subsurface and causing the warming observed in our Pacific BWT record. A decreasing SST gradient between the North and South Atlantic (Fig. 3D) further signifies the transfer of heat from the high northern latitudes. Thus, we argue that the expanded glaciation of Antarctica during the late Pliocene culminated in a fundamental change in deep water circulation, contributing to the intensification of NHG by facilitating a major redistribution of heat at Earth's surface.

## REFERENCES AND NOTES

1. H. J. Dowsett et al., *Sci. Rep.* **3**, 2013 (2013).
2. N. J. Shackleton, M. A. Hall, D. Pate, "Pliocene stable isotope stratigraphy of site 846," in *Proceedings of the Ocean Drilling Program, Scientific Results*, N. G. Pisias, L. A. Mayer, T. R. Janecek, A. Palmer-Julson, T. H. van Andel, Eds. (Ocean Drilling Program, College Station, TX, 1995), vol. 138, pp. 337–355.
3. L. E. Lisiecki, M. E. Raymo, *Paleoceanography* **20**, PA1003 (2005).
4. M. A. Maslin, G. H. Haug, M. Sarthein, R. Tiedemann, *Geol. Rundsch.* **85**, 452–465 (1996).
5. S. Steph et al., *Paleoceanography* **21**, PA4221 (2006).
6. T. D. Herbert, L. C. Peterson, K. T. Lawrence, Z. Liu, *Science* **328**, 1530–1534 (2010).
7. K. T. Lawrence, S. Sosdian, H. E. White, Y. Rosenthal, *Earth Planet. Sci. Lett.* **300**, 329–342 (2010).
8. A. Martinez-Garcia, A. Rosell-Melé, E. L. McClymont, R. Gersonde, G. H. Haug, *Science* **328**, 1550–1553 (2010).
9. S. Sosdian, Y. Rosenthal, *Science* **325**, 306–310 (2009).
10. T. M. Cronin, H. J. Dowsett, G. S. Dwyer, P. A. Baker, M. A. Chandler, *Mar. Micropaleontol.* **54**, 249–261 (2005).
11. G. Bartoli et al., *Earth Planet. Sci. Lett.* **237**, 33–44 (2005).
12. Supplementary materials, analytical methods, supporting text, and supplementary figures are available on Science Online. Original data is archived as a paleoclimate data set with the National Oceanic and Atmospheric Administration's National Climatic Data Center, available online at: [www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets](http://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets).
13. H. Elderfield et al., *Quat. Sci. Rev.* **29**, 160–169 (2010).
14. H. Elderfield, J. Yu, P. Anand, T. Kiefer, B. Nyland, *Earth Planet. Sci. Lett.* **250**, 633–649 (2006).
15. R. G. Fairbanks, R. K. Matthews, *Quat. Res.* **10**, 181–196 (1978).
16. M. S. Fanfille, D. J. DePaolo, *Geochim. Cosmochim. Acta* **70**, 3883–3904 (2006).
17. M. Mudelsee, M. E. Raymo, *Paleoceanography* **20**, PA4022 (2005).
18. H. Flesche Kleiven, E. Jansen, T. Fronval, T. M. Smith, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **184**, 213–223 (2002).
19. E. Jansen, T. Fronval, F. Rack, J. E. T. Channell, *Paleoceanography* **15**, 709–721 (2000).
20. J. Knies et al., *Earth Planet. Sci. Lett.* **387**, 132–144 (2014).
21. P. Fretwell et al., *The Cryosphere* **7**, 375–393 (2013).
22. J. L. Bamber et al., *The Cryosphere* **7**, 499–510 (2013).
23. S. Passchier, *Paleoceanography* **26**, PA4204 (2011).
24. C. P. Cook et al., *Nat. Geosci.* **6**, 765–769 (2013).



**Fig. 3. Evolution of  $\delta^{18}\text{O}_{\text{sw}}$  and other paleoclimate proxies across the late Pliocene intensification of NHG.** (A) Calculated  $\delta^{18}\text{O}_{\text{sw}}$  for Pliocene deep Pacific site 1208 (blue, this study) and the North Atlantic site 607 [red, (9)]. Shaded areas indicate 1 SE of filtered  $\delta^{18}\text{O}_{\text{sw}}$ . The dashed line indicates modern deep Pacific  $\delta^{18}\text{O}_{\text{sw}}$  ( $-0.15\text{‰}$ ). SMOW, standard mean ocean water. (B) Pacific minus Atlantic  $\delta^{18}\text{O}_{\text{sw}}$  showing a reduction in the interbasinal gradient at  $\sim 2.7$  Ma. (C) Southern Ocean opal mass accumulation rates (MAR) (ODP site 1096) (33). (D) Evolution of North (DSDP site 607 and ODP site 982) (7) and South Atlantic (ODP site 1090) (8) high-latitude SST gradients. Data were interpolated to even 2-ky spacing before differencing (thin lines); thick lines are 50-ky smoothed curves.



25. T. Naish *et al.*, *Nature* **458**, 322–328 (2009).
  26. R. McKay *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6423–6428 (2012).
  27. A. N. LeGrande, G. A. Schmidt, *Geophys. Res. Lett.* **33**, L12604 (2006).
  28. A. H. Orsi, G. C. Johnson, J. L. Bullister, *Prog. Oceanogr.* **43**, 55–109 (1999).
  29. Z. Zhang, K. H. Nisancioglu, U. S. Ninnemann, *Nat. Commun.* **4**, 1499 (2013).
  30. M. E. Raymo, B. Grant, M. Horowitz, G. H. Rau, *Mar. Micropaleontol.* **27**, 313–326 (1996).
  31. A. C. Ravelo, D. H. Andreasen, *Geophys. Res. Lett.* **27**, 1001–1004 (2000).
  32. J. D. Wright, K. G. Miller, *Paleoceanography* **11**, 157–170 (1996).
  33. C.-D. Hillenbrand, D. K. Fütterer, “Neogene to Quaternary deposition of opal on the continental rise west of the Antarctic Peninsula, ODP leg 178, sites 1095, 1096, and 1101,” *Proceedings of the Ocean Drilling Program, Scientific Results*, P. F. Barker, A. Camerlenghi, G. D. Acton, A. T. S. Ramsay, Eds. (Ocean Drilling Program, College Station, TX, 2001), vol. 178, pp. 1–33.
  34. D. M. Sigman, S. L. Jaccard, G. H. Haug, *Nature* **428**, 59–63 (2004).
  35. D. A. Hodell, K. A. Venz-Curtis, *Geochim. Geophys. Geosyst.* **7**, Q09001 (2006).
  36. A. Martínez-García *et al.*, *Nature* **476**, 312–315 (2011).
  37. L. Menviel, A. Timmermann, O. E. Timm, A. Mouchet, *Paleoceanography* **25**, PA4231 (2010).
  38. N. J. Shackleton, J. Imbrie, M. A. Hall, *Earth Planet. Sci. Lett.* **65**, 233–244 (1983).
- ACKNOWLEDGMENTS**  
We thank N. Venti and K. Billups for providing samples and J. Crisicone, N. Abdul, and R. Mortlock for lab assistance. Funding

was provided by NSF grants EAR-1052257 and OCE-1334691. B.K.C. thanks Rutgers Aresty Undergraduate Research Fellowship for its support.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6211/847/suppl/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S8  
References (39–84)

5 May 2014; accepted 10 October 2014  
Published online 23 October 2014;  
10.1126/science.1255586

## CLIMATE CHANGE

# Projected increase in lightning strikes in the United States due to global warming

David M. Roms, <sup>1\*</sup> Jacob T. Seeley, <sup>1</sup> David Vollaro, <sup>2</sup> John Molinari <sup>2</sup>

Lightning plays an important role in atmospheric chemistry and in the initiation of wildfires, but the impact of global warming on lightning rates is poorly constrained. Here we propose that the lightning flash rate is proportional to the convective available potential energy (CAPE) times the precipitation rate. Using observations, the product of CAPE and precipitation explains 77% of the variance in the time series of total cloud-to-ground lightning flashes over the contiguous United States (CONUS). Storms convert CAPE times precipitated water mass to discharged lightning energy with an efficiency of 1%. When this proxy is applied to 11 climate models, CONUS lightning strikes are predicted to increase  $12 \pm 5\%$  per degree Celsius of global warming and about 50% over this century.

Lightning exerts a powerful control on atmospheric chemistry through its generation of nitrogen oxides, especially in the middle and upper troposphere (1–3). As the primary trigger for wildfires, lightning also

shapes the evolution of species and ecosystems (4, 5). Despite its importance, the future increase in lightning flash rates due to global warming remains poorly constrained: Estimates range from 5% (6, 7) to over 100% (8) per degree Celsius (°C)

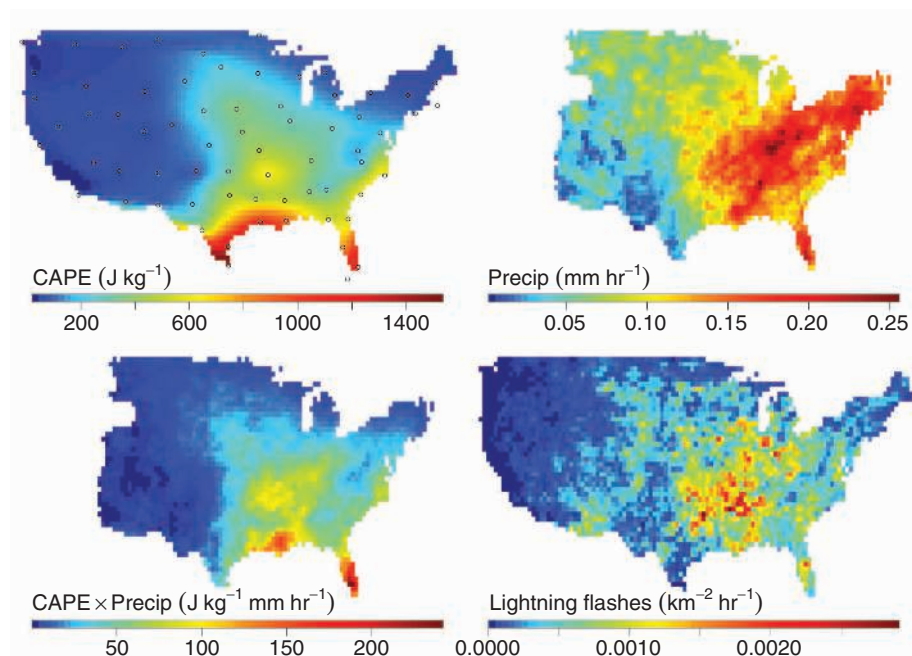
of global mean temperature increase. Here we show that a simple proxy—the product of the convective instability and the precipitation rate—explains most of the variance in lightning flashes over the contiguous United States (CONUS). When applied to global climate models (GCMs), this proxy predicts a mean increase in flash rate of 12% per global-mean °C over the CONUS. This augurs significant changes in the future atmospheric chemistry and wildfire frequency of North America.

Previous estimates of the sensitivity of lightning flash rates to global mean temperature have relied on ad hoc proxies for use in GCMs or observed correlations between temperature and lightning (or lightning proxies). One GCM parameterization sets the total flash rate (in flashes per minute) equal to a constant times the maximum cloud height to the fifth power (9), and later

<sup>1</sup>Department of Earth and Planetary Science, University of California, Berkeley, and Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

<sup>2</sup>Department of Atmospheric and Environmental Sciences, State University of New York at Albany, Albany, NY, USA.

\*Corresponding author. E-mail: romps@berkeley.edu



**Fig. 1. Mean maps of CAPE, precipitation, CAPE times precipitation, and lightning flashes.** For the year 2011, maps are shown of mean (top left) CAPE from the SPARC radiosonde data, (top right) precipitation from the National Weather Service River Forecast Center data, (bottom left) product of the top two maps, and (bottom right) CG lightning from the NLDN data. For CAPE, means are calculated by averaging all 00 and 12 GMT soundings; circles denote the locations of radiosonde releases. For precipitation and lightning, means are calculated by averaging over 22–02 and 10–14 GMT.

papers extended this to cloud-to-ground (CG) lightning [e.g., (10)]. Despite this proxy's lack of dependence on area (making it sensitive to GCM grid spacing) or on the amount of storm activity (making it insensitive to the rain rate), it is the most widely used proxy for lightning (6, 7, 11, 12), generating estimates for the increase in global flash rate ranging from 5% per °C to 16% per °C. Observational estimates also predict a wide range, but at much higher values. At some specific weather stations, flash rates increase with wet-bulb temperature over the seasonal cycle at rates in the range of 100 to 800% per °C (13). A similar analysis using satellite lightning data gives interannual sensitivities for Northern Hemisphere land and the globe of 56% per °C and 40% per °C, respectively (14). Using the fundamental mode of the Schumann resonance as a proxy for flash rate, a global sensitivity of 100% per °C is found (8). Although these estimates have the advantage of being rooted in observations, it is unlikely that these intraseasonal and interseasonal relationships are the same as those in a global warming scenario. By developing a new proxy for lightning that is physically motivated, testable with observations, and applicable to GCMs, a more robust projection is sought here for the CONUS, which is a major contributor to global lightning (15).

Here we propose that the lightning flash rate per area is proportional to the precipitation rate times convective available potential energy (CAPE). This proxy combines the observed linearity of flash rate on precipitation rate (16–19) with suggestions that flashes are positively correlated with CAPE (20–23). In mathematical form,

$$F = \frac{\eta}{E} \times P \times \text{CAPE} \tag{1}$$

where  $F$  is the lightning flash rate per area ( $\text{m}^{-2} \text{s}^{-1}$ ),  $P$  is the precipitation rate ( $\text{kg m}^{-2} \text{s}^{-1}$ ), and CAPE is in  $\text{J kg}^{-1}$ . Using an adiabatic definition of CAPE, the product of CAPE and  $P$  is the theoretical maximum rate at which kinetic energy is imparted to ascending water condensates, in units of  $\text{W m}^{-2}$ . The constant of proportionality,  $\eta/E$ , contains the dimensionless conversion efficiency  $\eta$  and the energy discharge per flash  $E$  (in joules). The efficiency  $\eta$  is the ratio of power per area dissipated by lightning to the CAPE per area per time available to condensates. We do not propose here a specific charging mechanism, but we note that most charging mechanisms are consistent with the notion that higher updraft speeds and water contents should yield higher flash rates.

In this study, we focused on the CONUS because it is well instrumented. Three sets of data were used, which overlap during the year 2011. CAPE is calculated from SPARC (Stratosphere-troposphere Processes And their Role in Climate) radiosonde data (24),  $P$  is taken from the National Oceanic and Atmospheric Administration River Forecast Centers (25), and CG lightning flashes are obtained from the National Lightning Detection Network [NLDN (26, 27)]; for more details, see the supplementary materials. The means of these quantities during 2011 are shown in Fig. 1.

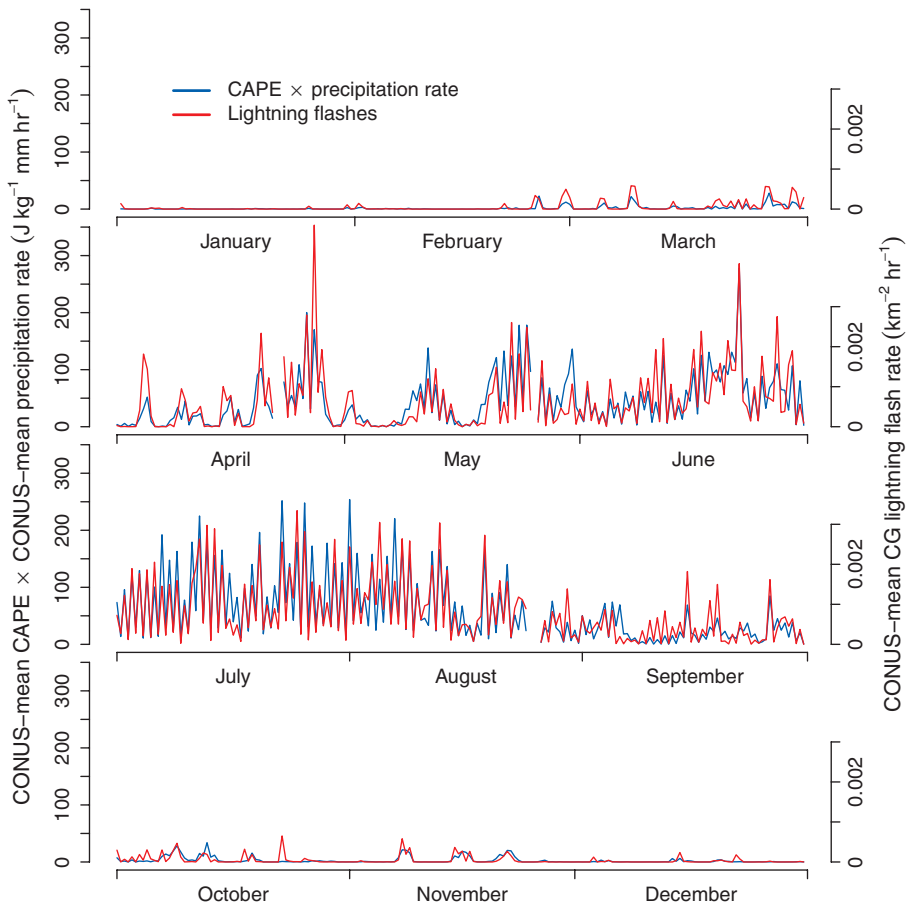
Also shown in Fig. 1 is the pointwise product of the annual mean CAPE map and the annual mean  $P$  map. This map of  $\text{CAPE} \times P$  (lower left panel) bears a close resemblance to the map of lightning flashes (lower right panel). In particular, both  $\text{CAPE} \times P$  and the flash rate maximize in Florida and in the states adjacent to the Mississippi and Ohio rivers. This

distribution of flashes is similar to the long-term annual mean [see Fig. 1 of (28)], because we are sampling lightning in the local morning and evening, which correspond to the trough and peak of the CONUS mean diurnal cycle [see Fig. 2 of (28)].

To assess the performance of the  $\text{CAPE} \times P$  proxy, we focus here on the time series of CONUS

**Table 1. Future changes predicted by GCMs.** Predicted changes in global mean temperature ( $\Delta T$ ) and percent per global mean °C changes in CONUS annual mean CAPE ( $\Delta \text{CAPE}$ ), precipitation ( $\Delta \text{Pr}$ ), and CG lightning flash rate ( $\Delta \text{CG}$ ) are shown for 11 CMIP5 GCMs. Changes are calculated for the years 2079–2088 of the RCP8.5 experiment relative to the years 1996–2005 of the historical experiment.

| GCM           | $\Delta T$ (°C) | $\Delta \text{CAPE}$ (%°C) | $\Delta \text{Pr}$ (%°C) | $\Delta \text{CG}$ (%°C) |
|---------------|-----------------|----------------------------|--------------------------|--------------------------|
| BCC-CSM1.1    | 3.4             | 6.4                        | -0.6                     | 3.4                      |
| BCC-CSM1.1(m) | 3.1             | 8.8                        | -0.2                     | 6.9                      |
| CanESM2       | 4.7             | 12.9                       | 4.2                      | 17.3                     |
| CCSM4         | 3.9             | 7.3                        | 2.0                      | 9.1                      |
| CNRM-CM5      | 3.9             | 9.9                        | 2.6                      | 12.2                     |
| FGOALS-g2     | 3.1             | 11.5                       | -1.8                     | 7.0                      |
| GFDL-CM3      | 5.0             | 16.5                       | 2.6                      | 17.6                     |
| GFDL-ESM2M    | 2.5             | 13.4                       | 2.7                      | 15.9                     |
| MIROC5        | 3.4             | 15.1                       | 0.3                      | 16.3                     |
| MRI-CGCM3     | 3.4             | 12.5                       | 3.0                      | 14.7                     |
| NorESM1-M     | 3.6             | 8.5                        | 1.4                      | 10.3                     |
| Mean:         | 3.6             | 11.2                       | 1.5                      | 11.9                     |



**Fig. 2. Time series of the proxy and observed lightning flashes.** For the year 2011, time series are shown of 0 and 12 GMT (blue) CONUS mean precipitation times CONUS mean CAPE and (red) CONUS mean CG lightning flash rate.

mean CAPE, precipitation, and lightning. A spatially resolved evaluation of Eq. 1 is not attempted here because of practical difficulties stemming from the sparsity of CAPE measurements and the fact that convection releases CAPE. The latter fact precludes using the product of colocated precipitation and CAPE; instead, a CAPE value nearby, but upwind, of a storm should be used. With the sparse network of noisy CAPE measurements, this is a challenging task that is left to future work.

Figure 2 plots two time series for the entire year of 2011: the product of CONUS mean CAPE and CONUS mean precipitation rate in blue (axes on the left), and CONUS mean CG lightning flash rate in red (axes on the right). The proxy varies in synchrony with the flash rate on a range of time scales, from the diurnal [high at 0 Greenwich mean time (GMT), low at 12 GMT] to the seasonal (high in spring and summer, low in fall and winter) and time scales in between (high during periods of sustained storminess, low in the lulls between). The proxy successfully captures the varying magnitude of the flash rate, as well. In the fall and winter, both the proxy and the flash rate have peak values about one order of magnitude smaller than their peak values in the spring and summer.

The lower right panel of Fig. 3 shows the scatter plot of the  $\text{CAPE} \times P$  and flash-rate time series. They are related to each other in a linear fashion, and the proxy explains 77% of the variance in the lightning flash rate. This can be compared with other candidates, such as  $P$ , the maximum height of convection [level of neutral buoyancy (LNB)] to the fifth power, and CAPE, which explain only 29, 39, and 52% of the variance, respectively. The  $\text{CAPE} \times P$  proxy also explains a substantial amount of the lightning variance within a season: 69% for January, February, March; 65% for April, May, June; 75% for July, August, September; and 40% for October, November, December.

Using the fact that 1 mm of precipitation equals  $1 \text{ kg m}^{-2}$  of liquid water, the best-fit line in the lower right panel is described by Eq. 1 with  $\eta/E = 1.3 \times 10^{-11} \text{ J}^{-1}$ . A best estimate for the energy released by a midlatitude CG lightning flash is about 1 GJ, although estimates range from a few tenths of a GJ to several GJ (*1*). Using  $E = 1 \text{ GJ}$ , we obtain an efficiency of  $\eta = 0.01$ . In other words, 1% of the CAPE that could be theoretically extracted by water (i.e., CAPE times the processed water mass) is converted to electrical potential energy that is then discharged by CG lightning.

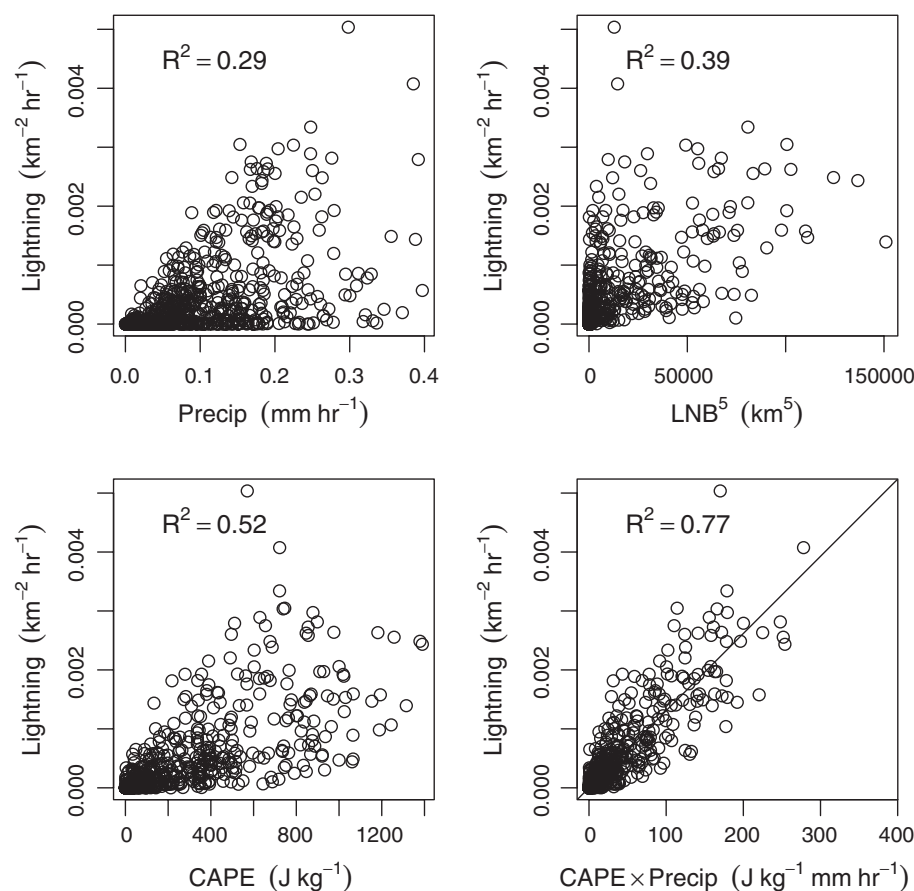
Given the success of  $\text{CAPE} \times P$  in replicating the time series of observed lightning flashes, it is a natural candidate for assessing future changes in flash rates due to global warming. Previous studies of GCM simulations have found that global warming causes CAPE to increase over much of Earth (29, 30) and over the CONUS in particular (31–33). Similar results have been found in much simpler cloud-resolving simulations, in which CAPE increases with sea surface temperature (34–36). Recent work has provided insight into why this increase in CAPE occurs (36). In addition, the global precipitation rate is expected to increase with global temperature (37), although the predicted changes in annual mean precipitation are of variable sign across the United States.

To assess how these future changes will affect lightning, we analyzed output from 11 GCMs in the Coupled Model Intercomparison Project Phase 5 [CMIP5 (38)]; see the supplementary materials for more details. Table 1 lists the CONUS mean fractional changes in CAPE and precipitation for the GCMs in our ensemble. To account for differences in climate sensitivity between the models, we report these results as percent changes per °C of global warming. All GCMs in our ensemble predict that CONUS mean CAPE will increase over the 21st century, with a mean increase of 11.2% per °C of global warming. There is a high level of agreement between the models on the spatial pattern and magnitude of this CAPE increase (fig. S1). On the other hand, there is significantly more variation in the GCMs' predictions for future precipitation; the mean response is a 1.5% increase per °C of global warming, but some models predict decreased precipitation over the CONUS.

The percent change in annual mean lightning flash rate can be estimated as the mean percent change in the product of CAPE and precipitation time series between the years 1996–2005 and 2079–2088. Using this method, all GCMs in our ensemble predict annual mean lightning-strike frequency in the United States to increase, with a mean increase of 12% per °C (column 5 of Table 1). The standard deviation of the ensemble's predictions is 5% per °C; therefore, we can conclude that the rate of CG lightning strikes over the CONUS is likely to increase as a function of global mean temperature at a rate of  $12 \pm 5\%$  per °C. Overall, the GCMs predict a ~50% increase in the rate of lightning strikes in the CONUS over the 21st century.

#### REFERENCES AND NOTES

1. U. Schumann, H. Huntrieser, *Atmos. Chem. Phys.* **7**, 3823–3907 (2007).
2. H. Levy II, W. J. Moxim, P. S. Kasibhatla, *J. Geophys. Res.* **101**, 22911 (1996).
3. V. Grewe, *Sci. Total Environ.* **374**, 167–181 (2007).
4. W. J. Bond, J. E. Keeley, *Trends Ecol. Evol.* **20**, 387–394 (2005).
5. J. G. Pausas, J. E. Keeley, *Bioscience* **59**, 593–601 (2009).
6. C. Price, D. Rind, *J. Geophys. Res.* **99**, 10823 (1994).
7. N. Michalon, A. Nassif, T. Saouri, J. F. Royer, C. A. Pontikis, *Geophys. Res. Lett.* **26**, 3097–3100 (1999).
8. E. R. Williams, *Science* **256**, 1184–1187 (1992).
9. C. Price, D. Rind, *J. Geophys. Res.* **97**, 9919–9933 (1992).



**Fig. 3. Lightning versus various proposed proxies.** For the year 2011, scatter plots are shown of the time series of 0 and 12 GMT CONUS mean lightning against (top left) precipitation, (top right) LNB to the fifth power, (bottom left) CAPE, and (bottom right) CAPE times precipitation.



10. C. Price, D. Rind, *Mon. Weather Rev.* **122**, 1930–1939 (1994).
11. A. Banerjee et al., *Atmos. Chem. Phys.* **14**, 9871–9881 (2014).
12. A. Krause, S. Kloster, S. Wilkenskild, H. Paeth, *J. Geophys. Res.: Biogeosci.* **119**, 312 (2014).
13. E. R. Williams, *Mon. Weather Rev.* **122**, 1917–1929 (1994).
14. N. Reeve, R. Tourni, Q. J. R. Meteorol. Soc. **125**, 893–903 (1999).
15. H. J. Christian et al., *J. Geophys. Res.* **108**, 4005 (2003).
16. L. J. Battan, *J. Atmos. Sci.* **22**, 79–84 (1965).
17. A. Tapia, J. A. Smith, M. Dixon, *J. Appl. Meteorol.* **37**, 1497–1509 (1998).
18. W. A. Petersen, S. A. Rutledge, *J. Geophys. Res.* **103**, 14025 (1998).
19. E. W. Meijer, P. F. J. van Velthoven, D. W. Brunner, H. Huntrieser, H. Kelder, *Phys. Chem. Earth, Part C Sol.-Terr. Planet. Sci.* **26**, 577–583 (2001).
20. E. R. Williams et al., *J. Atmos. Sci.* **49**, 1386–1395 (1992).
21. E. Williams et al., *J. Geophys. Res., D, Atmospheres* **107**, 8082 (2002).
22. S. D. Pawar, D. M. Lal, P. Murugavel, *Atmos. Res.* **106**, 44–49 (2012).
23. P. Murugavel, S. D. Pawar, V. Gopalakrishnan, *Int. J. Climatol.* **34**, 3179–3187 (2014).
24. SPARC, U.S. High Vertical Resolution Radiosonde Data, Stratosphere-troposphere Processes and their Role in Climate (2013).
25. D. Kitzmiller, D. Miller, R. Fulton, F. Ding, *J. Hydrol. Eng.* **18**, 133–142 (2013).
26. R. S. Wacker, R. E. Orville, *J. Geophys. Res.* **104**, 2151 (1999).
27. R. E. Orville, G. R. Huffines, *Mon. Weather Rev.* **129**, 1179–1193 (2001).
28. R. L. Holle, *Mon. Weather Rev.* **142**, 1037–1052 (2014).
29. B. Ye, A. D. Del Genio, K. K. W. Lo, *J. Clim.* **11**, 1997–2015 (1998).
30. A. H. Sobel, S. J. Camargo, *J. Clim.* **24**, 473–487 (2011).
31. R. J. Trapp et al., *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19719–19723 (2007).
32. R. J. Trapp, N. S. Diffenbaugh, A. Gluhovsky, *Geophys. Res. Lett.* **36**, L01703 (2009).
33. N. S. Diffenbaugh, M. Scherer, R. J. Trapp, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16361–16366 (2013).
34. D. M. Romps, *J. Atmos. Sci.* **68**, 123–138 (2011).
35. C. J. Muller, P. A. O’Gorman, L. E. Back, *J. Clim.* **24**, 2784–2800 (2011).
36. M. S. Singh, P. A. O’Gorman, *Geophys. Res. Lett.* **40**, 4398–4403 (2013).
37. I. M. Held, B. J. Soden, *J. Clim.* **19**, 5686–5699 (2006).
38. K. E. Taylor, R. J. Stouffer, G. A. Meehl, *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).

#### ACKNOWLEDGMENTS

D.M.R. acknowledges support from the Scientific Discovery through Advanced Computing (SciDAC) program funded by the U.S. Department of Energy Office of Advanced Scientific Computing Research and Office of Biological and Environmental Research, and by the U.S. Department of Energy’s Earth System Modeling, an Office of Science, Office of Biological and Environmental Research program under contract no. DE-AC02-05CH11231. J.T.S. acknowledges support from the National Science Foundation (NSF) Graduate Research Fellowship under grant no. DGE1106400. D.V. and J.M. acknowledge support from NSF under grant no. AGS132576. Thanks are due to the SPARC data center for archiving the high-resolution radiosonde data and making them publicly available. The authors are also grateful to J. Paul, B. Lawrence, K. Sugioaka, and N. Jeevanjee for their help with the precipitation data. Thanks also to three anonymous reviewers. Data sources are described in the online supplementary materials.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6211/851/suppl/DC1  
Methods  
Fig. S1  
References

23 July 2014; accepted 20 October 2014  
10.1126/science.1259100

#### CIRCADIAN RHYTHM

# Dysrhythmia in the suprachiasmatic nucleus inhibits memory processing

Fabian Fernandez,<sup>2</sup> Derek Lu,<sup>1</sup> Phong Ha,<sup>1</sup> Patricia Costacurta,<sup>1</sup> Renee Chavez,<sup>1</sup> H. Craig Heller,<sup>1</sup> Norman F. Ruby<sup>1\*</sup>

Chronic circadian dysfunction impairs declarative memory in humans but has little effect in common rodent models of arrhythmia caused by clock gene knockouts or surgical ablation of the suprachiasmatic nucleus (SCN). An important problem overlooked in these translational models is that human dysrhythmia occurs while SCN circuitry is genetically and neurologically intact. Siberian hamsters (*Phodopus sungorus*) are particularly well suited for translational studies because they can be made arrhythmic by a one-time photic treatment that severely impairs spatial and recognition memory. We found that once animals are made arrhythmic, subsequent SCN ablation completely rescues memory processing. These data suggest that the inhibitory effects of a malfunctioning SCN on cognition require preservation of circuitry between the SCN and downstream targets that are lost when these connections are severed.

Deficits in cognitive performance caused by disrupted circadian timing have become a growing concern among health care professionals (1). Recent clinical studies have found that age-related declines in circadian function can lead to mild cognitive impairment or dementia (2, 3). These memory deficits are not simply a consequence of poorer sleep, because reductions in circadian rhythm amplitude and robustness can accelerate progression of mild cognitive impairment or dementia even when sleep quality is maintained (2, 3). The observation that circadian timing is substantially weakened among people with Alzheimer’s disease has raised the possibility that cognitive deficits might be treated by improving circadian function (4–6). However, there are no mouse or rat models of adult-onset circadian dysfunction in genetically and neurologically intact animals living in standard laboratory conditions.

Rodent models of chronic circadian arrhythmia such as clock gene knockouts or surgical ablation of the central circadian pacemaker, the suprachiasmatic nucleus (SCN), exhibit no or only modest deficits in declarative memory. Clock gene knockouts of *cryptochrome 1 and 2* or *period 1 and 2* are arrhythmic, and they exhibit normal spontaneous alternation behavior, long-term spatial memory for food rewards, and contextual memory for environments associated with foot shock (7, 8), although *cryptochrome 1 and 2* knockouts do fail to learn time-place associations (7). *Bmal1* knockouts are also arrhythmic and exhibit normal contextual fear conditioning and novel object recognition; however, they navigate poorly in the Morris water maze (9). In the instances where knockout mice exhibit performance deficits, it is unclear whether memory impairments are due to arrhythmia, to pleiotropic

gene effects, or to abnormalities during brain development (10). Mice and rats with SCN lesions exhibit no substantial impairment in avoidance tasks, recognition memory, spatial learning, or reversal spatial learning but have modest deficits in contextual fear conditioning and in the water maze probe test (11–14). In some studies, SCN ablation actually improves task performance (11, 12).

The marginal effect of clock gene knockouts and SCN lesions on memory stands in stark contrast to the well-documented adverse effects of shift work and jet lag on cognition (15). One critical factor that often gets overlooked in relating animal circadian studies to human conditions is the fact that human dysrhythmia occurs while the SCN circuitry remains intact both genetically and structurally. We evaluated the possibility that intact SCN circuitry is necessary for circadian dysfunction to interfere with memory processing by using the Siberian hamster (*Phodopus sungorus*) model of circadian arrhythmia. This model has been used to study homeostatic sleep mechanisms, where its value compared to clock gene knockout and SCN lesion models of arrhythmia has been recognized (16).

Siberian hamsters exhibit phase-resetting responses to single light pulses that are typical of nocturnal rodents (Fig. 1A), yet their response to two light signals is quite different. When these animals are given a phase-advancing light signal on one night, followed by a phase-delaying light signal on the next night, circadian timing is completely abolished within a few days, even though each signal given alone does not disrupt their circadian organization (Fig. 1B). This disruptive phase shift (DPS) protocol causes arrhythmia by suppressing the amplitude of clock gene oscillations within the SCN to zero, thereby driving the clock to its singularity point (17, 18). Induction of arrhythmia at the genetic and behavioral levels occurs within a few days and lasts indefinitely despite the continued presence of a daily light-dark cycle (18, 19).

The DPS protocol allowed us to evaluate recognition and spatial memory in arrhythmic Siberian

<sup>1</sup>Biology Department, Stanford University, Stanford CA, USA.

<sup>2</sup>Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, CA, USA.

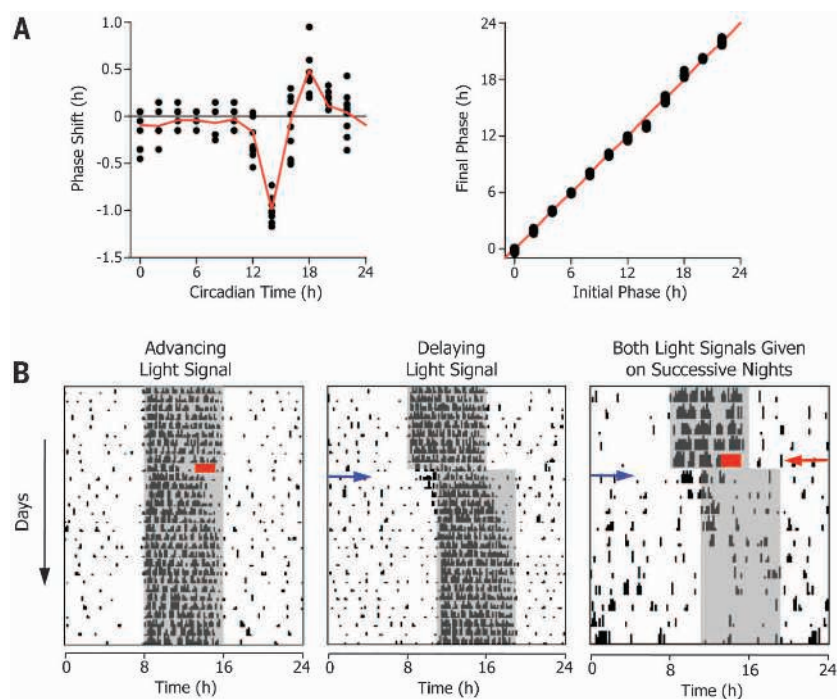
\*Corresponding author. E-mail: ruby@stanford.edu

hamsters with and without an intact SCN. Novel object recognition (NOR) was assessed as described (20), where the interval between the sample and test phases was 24 hours. Hamsters were assumed to have remembered the familiar object if they spend significantly more time investi-

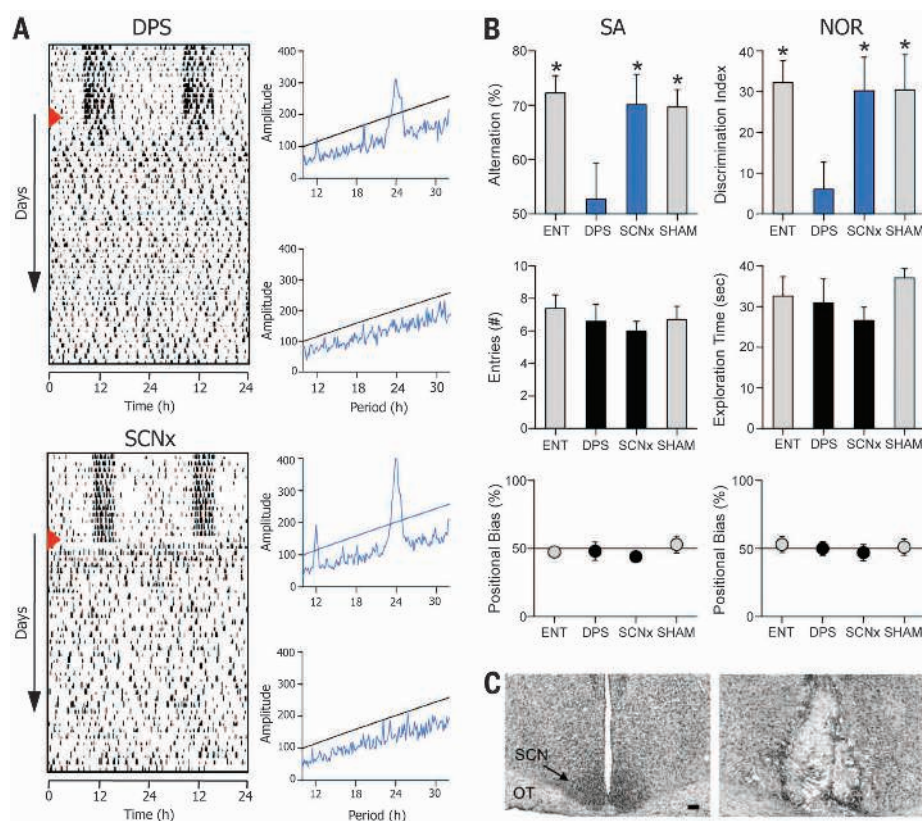
gating the novel object during the test phase. Spatial working memory was assessed by spontaneous alternation (SA) behavior in a T maze using a continuous trials procedure (27).

In the first experiment, adult Siberian hamsters were made arrhythmic either by the DPS

protocol or by bilateral surgical ablation of the SCN (Fig. 2C) (22). Arrhythmia was confirmed by chi-square periodogram analysis of locomotor activity rhythms in the animal's home cages during the 10 days before behavioral testing (Fig. 2A) (22). Lesions were confirmed by histological analysis



**Fig. 1. Phase-resetting responses of the Siberian hamster circadian system.** (A) Phase response (left) and transition (right) curves are typical for nocturnal rodents. Hamsters were housed in 16 hours of light followed by 8 hours of dark. Each animal was given a single light pulse (300 lux, 30 min) at different phases of the first circadian cycle in constant darkness (22). Each filled black circle represents a single animal ( $n = 98$ ); means of each time point are connected by the red line. The phase transition curve has a slope of 1, which indicates type 1 (i.e., weak) phase resetting (27). (B) The DPS protocol. Actograms show daily locomotor activity on each line, with successive days plotted from top to bottom. A phase-advancing signal (2-hour light pulse, red rectangle), followed by a phase-delaying signal (3-hour delay of the light-dark cycle, blue arrow) on the next day drives the system into arrhythmia. Night is indicated by the gray shaded area. The progressive daily shortening of the nightly active phase (i.e., alpha compression) after DPS treatment reliably predicts arrhythmia (17).



**Fig. 2. An intact SCN is necessary for circadian arrhythmia to impair recognition and spatial memory.** (A) Representative actograms of DPS (top) and SCNx (bottom) animals; red triangle indicates day of treatment. Chi-square periodograms confirm robust rhythms during entrainment (blue peaks above black line are significant;  $P = 0.001$ ) and arrhythmia. (B) SCNx animals ( $n = 11$ ) performed significantly better than random chance (i.e., 50 for SA, 0 for NOR), whereas DPS-treated animals ( $n = 8$ ) failed both memory tests (upper panels). ENT ( $n = 10$ ) and sham-operated (SHAM) ( $n = 11$ ) hamsters also performed significantly better than chance. Arrhythmic groups shown in blue. Scores significantly different from chance indicated by \* (one-sample  $t$  test;  $0.0001 < P < 0.01$ ). For DPS animals,  $P > 0.05$ . Number of T-maze arm entries and NOR exploration times did not differ significantly among ENT, DPS, SCNx, and SHAM groups ( $P > 0.05$ ; middle panels). No left-right positional biases were found in the T-maze or NOR arena ( $P > 0.05$ ; lower panels) (22). (C) Representative Nissl-stained tissue sections from intact (left) and SCNx (right) hamsters. Optic tracts are labeled OT. Black scale bar, 100  $\mu$ m.



of the SCN, which showed that sleep regulatory areas of the hypothalamus and the optic tract remained undamaged after the surgery (fig. S6). Hamsters were evaluated 4 to 6 weeks after surgery or after the DPS protocol. NOR testing was conducted over two consecutive days, followed by the SA test 3 days later. All tests were conducted within a 5-hour time window before dark onset in the colony room.

Consistent with studies done in mice and rats, hamsters made arrhythmic with complete bilateral ablation of the SCN (SCNx) performed just as well as normal entrained (ENT) animals in both NOR and SA tests (Fig. 2B and figs. S1A and S2A). By contrast, arrhythmic DPS hamsters did not perform better than chance in either test (Fig. 2B and figs. S1A and S2A). Surgery had no effect on the animal's engagement with the memory tests; exploration during the behavioral tests was the same for all groups (Fig. 2B and figs. S1A and S2A) (22). Because hippocampal lesions can cause perseverative responses in mazes (21), we also analyzed the data for left-right positional biases of SCNx hamsters in the open arena and the T maze, but none were found (Fig. 2B and figs. S1A and S2A) (22). We conclude that cognitive impairment caused by circadian dysfunction is predicated on maintaining circuit connections

from the SCN to other areas of the brain involved with information processing.

The idea that the arrhythmic SCN could actively inhibit memory formation was evaluated by a series of sequential treatments in a single group of animals. Siberian hamsters were first tested for NOR and SA while they were entrained to the light-dark cycle to establish baseline performance and then subjected to the DPS protocol. Once arrhythmia was confirmed (Fig. 3A), they were tested a second time 4 to 6 weeks later and then subjected to SCN ablation or sham surgery. After 4 to 6 weeks of postsurgical recovery, these animals were tested a third and final time. SCN lesions sometimes reduce overall levels of locomotor activity (23). Because such reductions could potentially limit exploration time in the memory tests, we examined the data for any changes in locomotor activity during the various treatment phases, but none were found (one-way analysis of variance;  $P > 0.05$ ) (Fig. 3B).

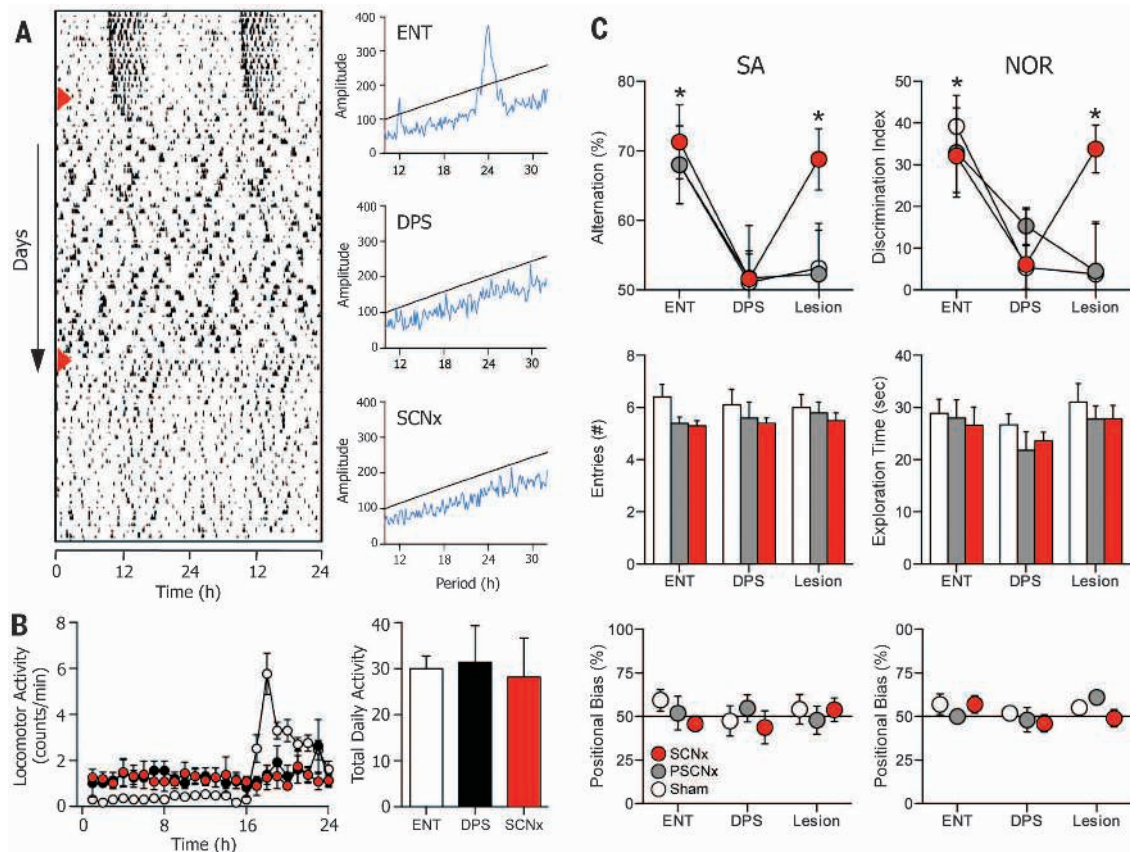
Animals successfully mastered the NOR and SA tests when they were entrained, but arrhythmia induced by the DPS protocol impaired their recognition and spatial memory (Fig. 3C and figs. S1B and S2B) (22). These memory deficits were completely reversed in hamsters that subsequently underwent SCN ablation surgery (Fig. 3C and

figs. S1B and S2B). Bilateral ablation of the SCN improved test performance to levels observed during entrainment. Exploratory activity and positional responses during the tests remained unaffected by the DPS protocol or by SCN surgery (supplementary text and movies S1 to S3), and sham-operated animals did not improve their performance (Fig. 3C and figs. S1B and S2B). We also found that memory was not restored in hamsters that sustained only partial ablation of the SCN (PSCNx) (Fig. 3C and figs. S1B and S2B). In PSCNx hamsters, the extent of the ablation was restricted between the rostral and middle SCN and damaged only about 20 to 60% of the nucleus.

The present study demonstrates that chronic arrhythmia per se does not cause memory impairments in animals. Rather, an arrhythmic SCN is necessary to realize these deficits, even if only a subpopulation of SCN neurons remains, as was the case for PSCNx animals. Lesion studies have consistently shown that circadian rhythms persist even if as little as 10% of SCN neurons survive ablation surgery (24) (supplementary text). This also appears to be the case for SCN control of memory processing. A number of studies have proposed that cognitive impairments caused by circadian dysfunction in humans are due to SCN cell loss or degeneration (5), but data from

**Fig. 3. SCN ablation rescues memory in arrhythmic animals.**

(A) Representative actogram of a hamster during ENT, DPS, and SCNx phases of the experiment. Phases were separated by 4 to 6 weeks. Days of DPS and SCNx treatments are indicated (red arrowheads). Periodograms evaluated the last 10 days of data from each phase of the experiment. (B) Locomotor activity was redistributed across the 24-hour day when animals became arrhythmic (left), but total daily activity is conserved across all conditions (right). (C) Performance on both memory tests (upper panels) was normal while animals were entrained but was impaired after the DPS protocol made them arrhythmic. SCN ablation subsequently rescued test performance (Lesion;  $n = 11$ ; red circles). Scores significantly different from chance indicated by \* (one-sample  $t$  test;  $0.0001 < P < 0.05$ ). Memory did not improve in SHAM animals (Lesion;  $n = 8$ ; white circles) or in animals with partial damage to the SCN (Lesion;  $n = 9$ ; gray circles). Surgery did not alter the number of arm entries or exploration time (middle panels) among SCNx (red bars), PSCNx (gray bars), or SHAM (white bars) animals (22). (C) None of the treatment groups showed any positional bias in the T maze or NOR arena (lower panels) (22).





PSCNx hamsters suggest that cell loss is not a direct cause of these impairments. Instead, gradual cell loss in humans might lead to changes in the remaining SCN circuitry that actively restrict the brain's ability to form new memory representations (25). The rescue of memory by SCN ablation seems to be due to a direct effect of SCN ablation rather than to indirect effects on sleep or damage to extra-SCN regions (figs. S3 to S6 and supplementary text).

These findings might bridge the gap between mechanistic neurobiological research in rodents and the understanding of human disorders of dysrhythmia because they open up the possibility that SCN connections to the medial temporal lobe, among other sites, might be targeted for pharmacotherapeutic treatments of cognitive impairment. Based on our past work showing that a  $\gamma$ -aminobutyric acid type A (GABA<sub>A</sub>) antagonist can rescue memory in DPS arrhythmic hamsters (19, 20), we proposed that an arrhythmic SCN might produce a steady GABAergic signal to the septal nuclei that attenuates cholinergic activation of the hippocampus, and septal activation is necessary for normal SA behavior (19). Thus, the present study supports the idea that the SCN influences memory through an inhibitory mechanism that impairs memory processing when SCN function is compromised.

Circadian dysregulation facilitates progression of clinically diagnosed mild cognitive impairment and Alzheimer's disease (4–6). Although counterintuitive, pharmacotherapies designed around silencing the SCN might also be applicable in treating memory problems associated with these disorders; however, potential adverse consequences of this approach must be carefully weighed (26).

## REFERENCES AND NOTES

- I. Rouch, P. Wild, D. Ansiau, J. C. Marquié, *Ergonomics* **48**, 1282–1293 (2005).
- G. J. Tranah et al., *Ann. Neurol.* **70**, 722–732 (2011).
- G. E. Schlosser Covell et al., *Neurologist* **18**, 426–429 (2012).
- E. J. W. van Someren et al., *Biol. Psychiatry* **40**, 259–270 (1996).
- A. N. Coogan et al., *Biol. Psychiatry* **74**, 333–339 (2013).
- Y.-H. Wu, D. F. Swaab, *Sleep Med.* **8**, 623–636 (2007).
- E. A. Van der Zee et al., *Curr. Biol.* **18**, 844–848 (2008).
- C. Mulder, E. A. Van Der Zee, R. A. Hut, M. P. Gerkema, *J. Biol. Rhythms* **28**, 367–379 (2013).
- S. M. Wardlaw, T. X. Phan, A. Saraf, X. Chen, D. R. Storm, *Learn. Mem.* **21**, 417–423 (2014).
- E. A. Yu, D. R. Weaver, *Aging* **3**, 479–493 (2011).
- F. K. Stephan, N. S. Kovacevic, *Behav. Biol.* **22**, 456–462 (1978).
- R. E. Mistlberger, M. H. de Groot, J. M. Bossert, E. G. Marchant, *Brain Res.* **739**, 12–18 (1996).
- S. W. Cain, M. R. Ralph, *Neurobiol. Learn. Mem.* **91**, 81–84 (2009).
- T. X. Phan, G. C. Chan, C. B. Sindreu, K. L. Eckel-Mahan, D. R. Storm, *J. Neurosci.* **31**, 10640–10647 (2011).
- B. L. Smarr, K. J. Jennings, J. R. Driscoll, L. J. Kriegsfeld, *Behav. Neurosci.* **128**, 283–303 (2014).
- T. Deboer, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **287**, R8–R9 (2004).
- S. Steinlechner, A. Stieglitz, T. Ruf, *J. Biol. Rhythms* **17**, 248–258 (2002).
- B. P. Grone et al., *J. Biol. Rhythms* **26**, 78–81 (2011).
- N. F. Ruby et al., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 15593–15598 (2008).
- N. F. Ruby et al., *PLOS ONE* **8**, e72433 (2013).
- R. M. Deacon, J. N. Rawlins, *Nat. Protoc.* **1**, 7–12 (2006).
- Materials and methods are available as supplementary materials on Science Online.
- C. H. Ko, R. J. McDonald, M. R. Ralph, *Biol. Rhythm Res.* **34**, 177–192 (2003).
- N. F. Ruby, J. Dark, D. E. Burns, H. C. Heller, I. Zucker, *J. Neurosci.* **22**, 357–364 (2002).
- D. G. Harper et al., *Brain* **131**, 1609–1617 (2008).
- F. W. Turek, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **294**, R1672–R1674 (2008).
- A. T. Winfree, *The Geometry of Biological Time* (Springer-Verlag, New York, 1980).

## ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute of Mental Health (#MH095837).

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6211/854/suppl/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S6  
Movies S1 to S3  
References (28–34)

5 August 2014; accepted 17 October 2014  
10.1126/science.1259652

## CILIA AND FLAGELLA

# A molecular ruler determines the repeat length in eukaryotic cilia and flagella

Toshiyuki Oda,<sup>1\*</sup> Haruaki Yanagisawa,<sup>1\*</sup> Ritsu Kamiya,<sup>2,3</sup> Masahide Kikkawa<sup>1†</sup>

Existence of cellular structures with specific size raises a fundamental question in biology: How do cells measure length? One conceptual answer to this question is by a molecular ruler, but examples of such rulers in eukaryotes are lacking. In this work, we identified a molecular ruler in eukaryotic cilia and flagella. Using cryo-electron tomography, we found that FAP59 and FAP172 form a 96-nanometer (nm)–long complex in *Chlamydomonas* flagella and that the absence of the complex disrupted 96-nm repeats of axonemes. Furthermore, lengthening of the FAP59/172 complex by domain duplication resulted in extension of the repeats up to 128 nm, as well as duplication of specific axonemal components. Thus, the FAP59/172 complex is the molecular ruler that determines the 96-nm repeat length and arrangements of components in cilia and flagella.

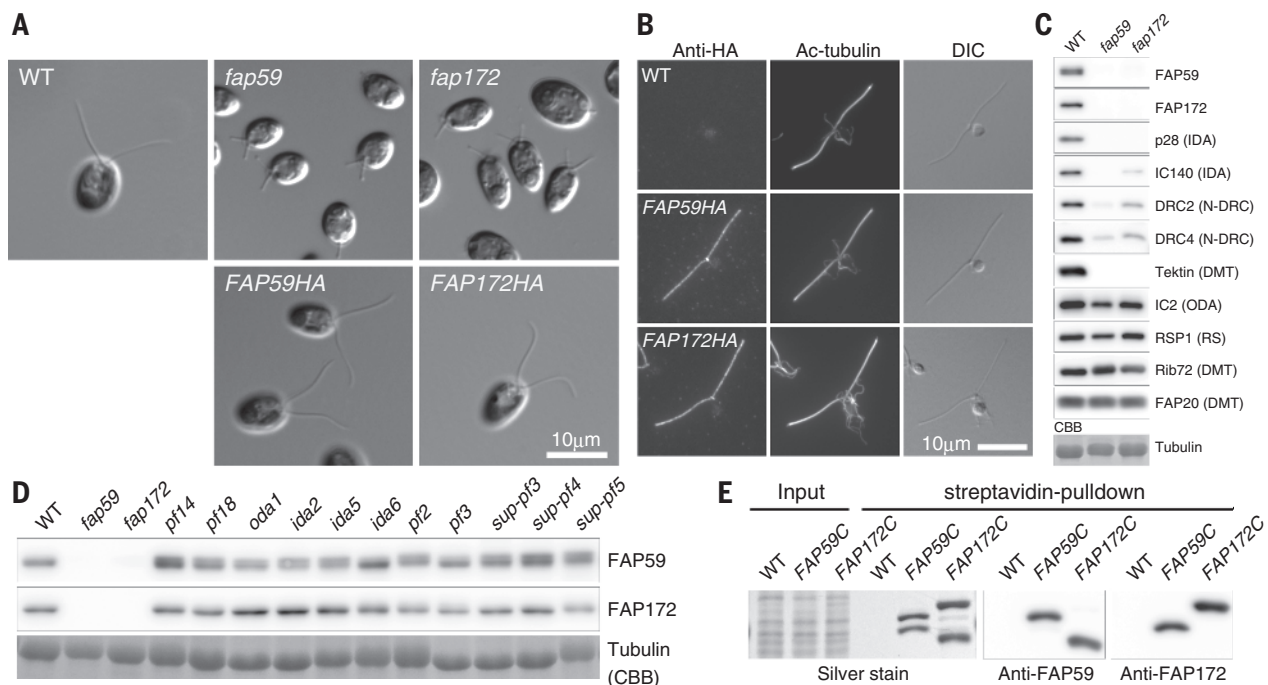
Cilia and flagella are motile organelles that play crucial roles in the generation of fluid flow by beating motion and in the development of mammals (1, 2). Their doublet structure has 96-nm repeats, which are very accurately specified. One explanation for this precise length control is that a molecular ruler exists (3). In bacteria and bacteriophages, ruler molecules have been shown to determine the length of the injectosome needles and the tail structures (4, 5). In eukaryotes, however, titins and nebulins have been proposed to be the rulers for muscle (6, 7), but this has been controversial (8–10).

To investigate the mechanism of ciliary and flagellar assembly, we focused on two *Chlamydomonas* proteins, FAP59 and FAP172. A lack of these proteins results in axonemal disorganization in various mammals and zebrafish (11–14). We first isolated a FAP172-missing mutant (strain

5-8) by screening *Chlamydomonas* motility mutants that lack FAP59 and FAP172 (table S1 and fig. S1, A and B; also see supplementary materials and methods). The mutant produced short and immotile flagella. Another four mutants that exhibit similar phenotypes to strain 5-8 were identified, including FAP172-deficient (*pf7*) or FAP59-deficient (*pf8*, *HA24*, and *gam5*) mutants (Fig. 1A, table S1, and fig. S1, B and C). Hereafter, we labeled *pf8* as *fap59* and *pf7* as *fap172*. Transformation of hemagglutinin (HA)-tagged genes to *fap59* and *fap172* mutants recovered wild-type (WT) motility and flagella length, confirming the direct relation between mutations in the two genes and the phenotypes (Fig. 1B, table S1, and fig. S1D). In *fap59* and *fap172* flagella, inner dynein arm (IDA) and nexin-dynein regulatory complex (N-DRC) components were absent or severely reduced (Fig. 1C). On the other hand, FAP59 and FAP172 were retained by mutant strains deficient in major axonemal components (Fig. 1D and table S2). Biochemical characterization revealed that FAP59 and FAP172 are mainly localized in flagella and released by treatment with 0.2 to ~0.4 M NaI and that FAP59 protein is phosphorylated in flagella (fig. S1, E to G) (14, 15).

Next we examined whether FAP59 and FAP172 form a complex because the flagellar localization of each protein depends on the other (Fig. 1C) (13).

<sup>1</sup>Department of Cell Biology and Anatomy, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo, 113-0033, Japan. <sup>2</sup>Department of Biological Science, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo, 113-0033, Japan. <sup>3</sup>Department of Life Science, Faculty of Science, Gakushuin University, 1-5-1 Meiji, Toshima-ku, Tokyo, 171-8588, Japan.  
\*These authors contributed equally to this work. †Corresponding author. E-mail: mkikkawa@m.u-tokyo.ac.jp



**Fig. 1. Biochemical characterization of *fap59* and *fap172* mutants.** (A) Differential interference contrast (DIC) images of *Chlamydomonas* cells. Expression of HA-tagged FAP59 and FAP172 rescued *fap59* and *fap172* mutants, respectively, is shown. (B) Immunofluorescence of WT and rescued nucleoflagellar apparatuses. Immunostaining of HA-tagged FAP59 and FAP172 indicates that the two proteins are localized along the length of the flagella. (C) Immunoblots of axonemal proteins from WT, *fap59*, and *fap172* cells. Flagellar localizations of FAP59 and FAP172 are dependent on each other. Components of the IDA (p28 and IC140) and N-DRC (DRC2 and DRC4) are missing or severely reduced in

*fap59* and *fap172* mutants. Tektin, a DMT-associated protein, is also missing in the mutants. IC2, an intermediate chain of the ODAs, and RSP1, a component of the RSs, appear unchanged. Components of the DMT (Rib72 and FAP20) are not reduced, either. (D) Immunoblot analysis of axonemal proteins. Except for *fap59* and *fap172*, all of the mutants retain FAP59 and FAP172 proteins. See table S2 for descriptions of the mutant strains used in this study. CBB, Coomassie Brilliant Blue. (E) Pull-down assay. BCCP tags were added to the C termini of FAP59 or FAP172, and the BCCP-tagged proteins were pulled-down by streptavidin beads. FAP172 was copurified with BCCP-tagged FAP59 and vice versa.

We labeled the two proteins with biotin carboxyl carrier protein (BCCP) tags and performed pull-down assays using streptavidin beads (Fig. 1E). BCCP-tagged FAP59 pulled down FAP172 and vice versa, showing that FAP59 and FAP172 form a complex in flagella.

To examine the structural defects caused by a loss of the FAP59/172 complex, we used cryo-electron tomography (cryo-ET) to observe the mutant axonemes. In WT axonemes, the doublet microtubule (DMT) components are arranged with 96-nm periodicity, as represented by radial spokes (RSs) (Fig. 2, A and B) (16, 17). In *fap59* and *fap172*, however, RSs were attached to DMTs with an irregular periodicity of ~32 nm instead of 96 nm (Fig. 2, C and D, and fig. S2, A and B). In agreement with our immunoblot analyses (Fig. 1C) and previous reports (12, 13), IDAs and the N-DRC were missing in the mutant axonemes (fig. S2C and fig. S3, A, B, and G to L). Thus, the FAP59/172 complex is essential for the establishment of 96-nm repeats in DMTs. The 24-nm repeats of outer dynein arms (ODAs) were unaltered in these mutants (fig. S3, C to E), suggesting that the arrangement of ODAs does not depend on the FAP59/172 complex. Therefore, we hypothesized that the FAP59/172 complex serves as a molecular ruler for the 96-nm repeats of DMTs.

To test an alternative hypothesis that tektin is a 96-nm molecular ruler (18, 19), we also exam-

ined a N-DRC mutant, *sup-pf-5* (20, 21). Tektin is missing in *sup-pf-5* flagella (Fig. 2F), but its axonemes retain the 96-nm periodicity of RSs (Fig. 2E). Thus, tektin does not define the 96-nm repeats of DMTs.

To examine the possible role of the FAP59/172 complex as a molecular ruler, we inserted BCCP tags into the two proteins at different sites and determined their three-dimensional (3D) positions on DMTs using cryo-ET and an enhanced streptavidin-labeling method (Fig. 3 and fig. S4) (22, 23). The labels on FAP59 and FAP172 were similarly positioned along the protofilament(s), and each label appeared once per 96-nm repeat (Fig. 3, B and C). The labels in the middle segment were located approximately halfway between the N and C termini. Thus, both FAP59 and FAP172 form 96-nm-long structures and lie along the long axes of DMTs, consistent with the idea that the FAP59/172 complex is a 96-nm molecular ruler.

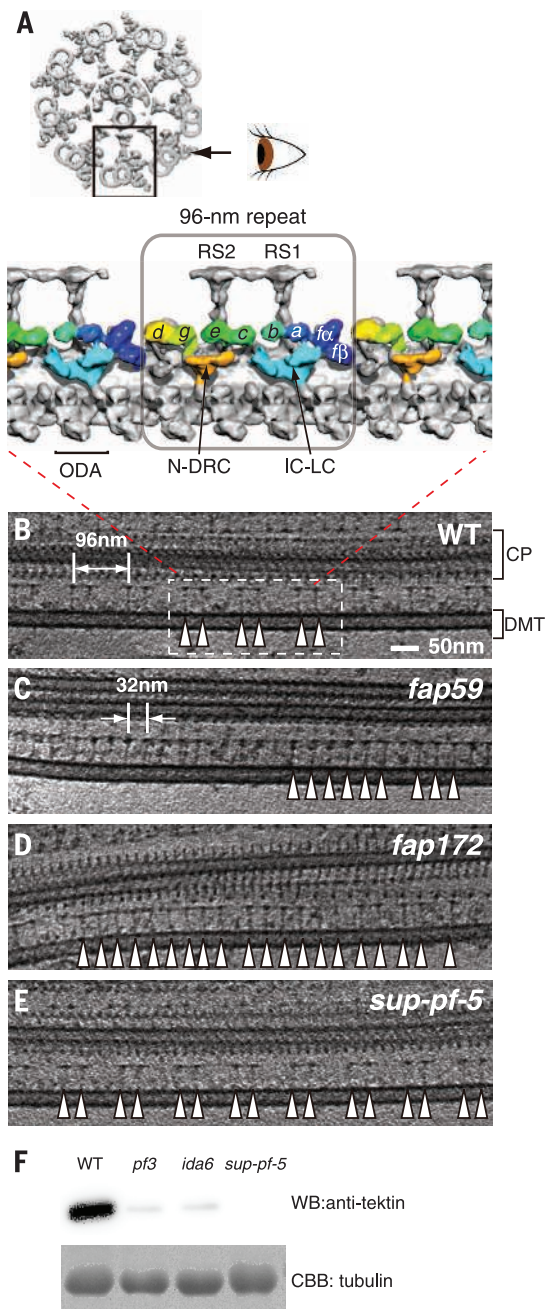
To test whether the FAP59/172 complex defines the 96-nm DMT repeats, we expressed longer versions of the complex in *fap59 fap172* double mutants and used cryo-ET to observe their axonemal ultrastructures (Fig. 4). The FAP59/172 complex was lengthened by duplicating the coiled-coil domains at the N terminal, at the middle, or near the C-terminal part. The short flagella phenotype of the double mutant was rescued by expressing the elongated proteins (fig. S5, A to E).

Cryo-ET of the axonemes of the rescued strains—named *FAP59/172-EL1*, *-EL2*, and *-EL3*—revealed that the repeat length was extended to ~128 nm (*FAP59/172-EL1* and *-EL3*) or ~120 nm (*FAP59/172-EL2*) (Fig. 4, C to H). Extension of the repeats by 32 nm in *FAP59/172-EL1* and *-EL3*, and by 24 nm in *FAP59/172-EL2*, is approximately proportional to the number of amino acids duplicated in each mutant (~1.3 Å per residue).

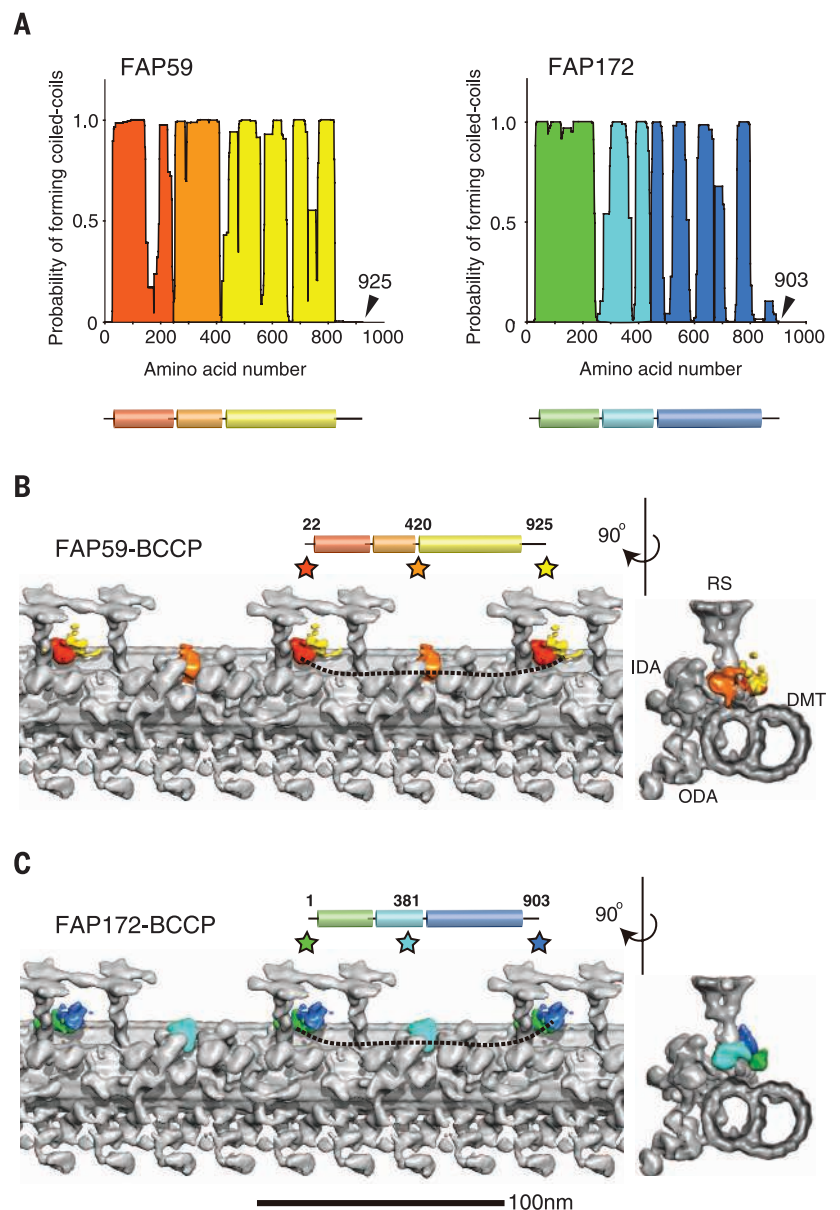
Elongation of the repeats was accompanied by duplication of specific structures in the flagella. In *FAP59/172-EL1*, the IDA *f* (Fig. 4D, deep blue), IDA *a* (blue), the intermediate chain-light chain (IC-LC) complex (light blue), and RS1 (RS1') were duplicated. In *FAP59/172-EL2*, the gap between IDAs *f* and *d* was widened (Fig. 4F, square brackets), and weak densities of additional IDA-like structures (purple) were observed. In *FAP59/172-EL3*, the single-headed IDAs *c* and *e* (Fig. 4H, green), N-DRC (orange), and RS2 (RS2') were duplicated. These results strongly support the idea that each domain of the FAP59/FAP172 complex recruits specific axonemal components to the DMTs.

Taken together, our results indicate that the FAP59/FAP172 complex uses very intricate mechanisms to construct the 96-nm repeat of DMT, rather than simply serving as a physical ruler. We propose “positive” and “negative” scaffolding mechanisms for the correct alignment of





**Fig. 2. The 96-nm periodicity of radial spokes was lost in *fap59* and *fap172* axonemes.** (A) The 96-nm repeating units of DMTs in a *Chlamydomonas* WT axoneme. Each 96-nm repeat has two RSs (RS1 and RS2), seven IDAs (a to e,  $\alpha$ ,  $\beta$ , g, and an IC-LC complex of f), four ODAs, and one N-DRC. (B to E) Slices of tomograms showing positions of the RSs (white arrowheads). (B) The WT axoneme shows the regular 96-nm periodicity of the RSs. Distal ends are to the left. CP, central pair microtubules. (C and D) Both *fap59* and *fap172* axonemes show irregular alignment of the RSs with ~32-nm gaps. (E and F) *sup-pf-5* axonemes show a normal 96-nm repeat, in spite of tektin protein deficiency. (F) Immunoblot analysis of N-DRC mutant axonemes. An absence of tektin proteins in the *sup-pf-5* mutant and a reduction of tektin protein in the *pf3* and *ida6* mutants were observed. WB, Western blot.



**Fig. 3. Both FAP59 and FAP172 take on a 96-nm-long extended conformation.** (A) Coiled-coil structure prediction of FAP59 and FAP172 using the CoilScan program (24). The predicted coiled-coil domains were roughly divided into three parts. (B and C) FAP59 and FAP172 were structurally tagged, and their positions were visualized by comparing the averaged subtomograms of the WT DMT (gray) with those of the labeled DMTs. The positions of tags (indicated by stars) are shown above each tomogram. Views on the right were rotated 90° around the vertical axis. (B) BCCP tags were added to the N terminus (red, after Met<sup>22</sup>) and the C terminus (yellow, before the stop codon) or inserted into the middle (orange, after Asp<sup>420</sup>) of FAP59. (C) BCCP tags were added to the N terminus (green, before Met<sup>1</sup>) and the C terminus (blue, before the stop codon) or inserted into the middle (light blue, after Leu<sup>381</sup>) of FAP172. Distal ends are to the left. The orientation of the proteins relative to the DMTs was determined in combination with the next experiment (Fig. 4).



RSs, IDAs, and N-DRCs (fig. S6). In the absence of the FAP59/172 complex, an excessive number of RSs bind along the specific protofilament(s) of the A-tubule, indicating that binding of the RSs to the A-tubule does not depend on the FAP59/172 complex. Therefore, the FAP59/172 complex can be seen as a kind of negative regulator for RSs, as the complex masks most of the RS-binding regions. At the same time, the complex leaves “holes” that allow for binding of RSs in appropriate locations. In contrast, the absence of the FAP59/172 complex causes loss of IDAs and the N-DRC, suggesting that the complex works as a positive regulator and provides anchoring sites for IDAs and N-DRCs.

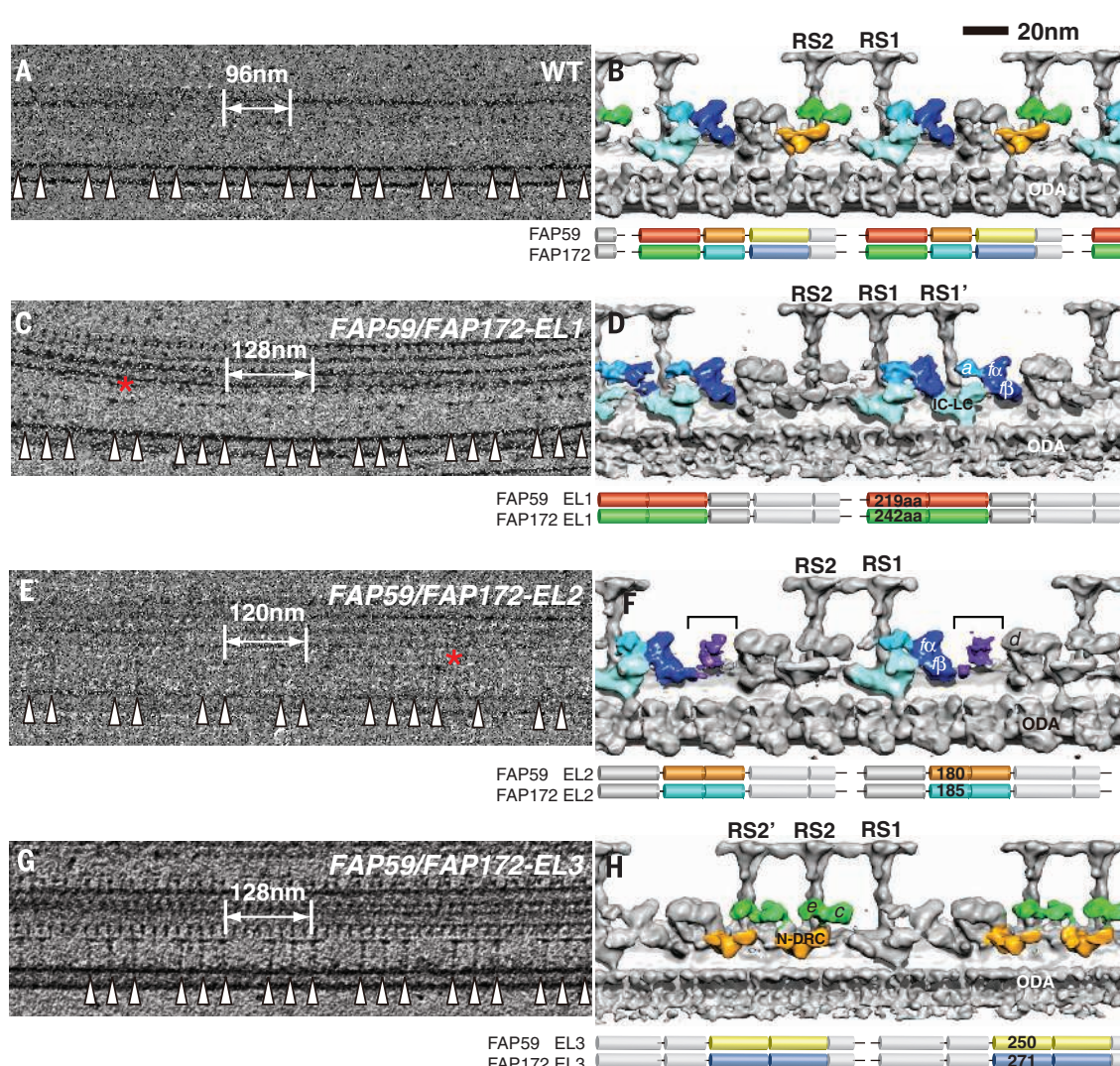
1. L. T. Haimo, J. L. Rosenbaum, *J. Cell Biol.* **91**, 125s–130s (1981).
2. G. J. Pazour, J. L. Rosenbaum, *Trends Cell Biol.* **12**, 551–555 (2002).
3. G. R. Cornelis, C. Agrain, I. Sorg, *Curr. Opin. Microbiol.* **9**, 201–206 (2006).

3. I. Katsura, *Nature* **327**, 73–75 (1987).
  5. L. Journet, C. Agrain, P. Broz, G. R. Cornelis, *Science* **302**, 1757–1760 (2003).
  6. J. Trinick, *Trends Biochem. Sci.* **19**, 405–409 (1994).
  7. M. Kruger, J. Wright, K. Wang, *J. Cell Biol.* **115**, 97–107 (1991).
  8. I. Fernandes, F. Schöck, *J. Cell Biol.* **206**, 559–572 (2014).
  9. M. L. Greaser, J. M. Pleitner, *Front. Physiol.* **5**, 35 (2014).
  10. A. Castillo, R. Nowak, K. P. Littlefield, V. M. Fowler, R. S. Littlefield, *Biophys. J.* **96**, 1856–1865 (2009).
  11. T. S. McClintock, C. E. Glasser, S. C. Bose, D. A. Bergman, *Physiol. Genomics* **32**, 198–206 (2008).
  12. A. C. Merville et al., *J. Vet. Intern. Med.* **28**, 771–778 (2014).
  13. A. Becker-Heck et al., *Nat. Genet.* **43**, 79–84 (2011).
  14. J. Lin et al., *J. Biol. Chem.* **286**, 29175–29191 (2011).
  15. J. Boesger, V. Wagner, W. Weisheit, M. Mittag, *Int. J. Plant Genomics* **2012**, 581460 (2012).
  16. K. H. Bui, T. Yagi, R. Yamamoto, R. Kamiya, T. Ishikawa, *J. Cell Biol.* **198**, 913–925 (2012).
  17. T. Heuser, M. Raytchev, J. Krell, M. E. Porter, D. Nicastro, *J. Cell Biol.* **187**, 921–933 (2009).
  18. J. M. Norrander, C. A. Perrone, L. A. Amos, R. W. Linck, *J. Mol. Biol.* **257**, 385–397 (1996).
  19. S. M. King, *Cell Motil. Cytoskeleton* **63**, 245–253 (2006).
  20. B. Huang, Z. Ramanis, D. J. Luck, *Cell* **28**, 115–124 (1982).
  21. G. Piperno, K. Mead, W. Shestak, *J. Cell Biol.* **118**, 1455–1463 (1992).
  22. T. Oda, M. Kikkawa, *J. Struct. Biol.* **183**, 305–311 (2013).
  23. T. Oda, H. Yanagisawa, T. Yagi, M. Kikkawa, *J. Cell Biol.* **204**, 807–819 (2014).
  24. A. Lupas, M. Van Dyke, J. Stock, *Science* **252**, 1162–1164 (1991).
- ## ACKNOWLEDGMENTS
- We thank M. Hirono (University of Tokyo) and T. Yagi (Prefectural University of Hiroshima) for technical assistance and valuable advice. This work was supported by CREST, Japan Science and Technology Agency (to M.K.), the Takeda Science Foundation (to M.K. and T.O.), the Kazato Research Foundation (to T.O.), and Japan Society for the Promotion of Science KAKENHI grant 20770119 (to H.Y.). We declare no competing financial interests. The electron microscopy maps of averaged DMT are available in the EM Data Bank ([www.emdatabank.org](http://www.emdatabank.org)) under accession numbers EMD-6108 to EMD-6117.
- ## SUPPLEMENTARY MATERIALS
- [www.sciencemag.org/content/346/6211/857/suppl/DC1](http://www.sciencemag.org/content/346/6211/857/suppl/DC1)  
Materials and Methods  
Supplementary Text  
Figs. S1 to S6  
Tables S1 to S3  
References (25–59)
- 20 August 2014; accepted 14 October 2014  
10.1126/science.1260214

We thank M. Hiroho (University of Tokyo) and T. Yagi (Prefectural University of Hiroshima) for technical assistance and valuable advice. This work was supported by CREST, Japan Science and Technology Agency (to M.K.), the Takeda Science Foundation (to M.K. and T.O.), the Kazato Research Foundation (to T.O.), and Japan Society for the Promotion of Science KAKENHI grant 20770119 (to H.Y.). We declare no competing financial interests. The electron microscopy maps of averaged DMT are available in the EM Data Bank ([www.emdatabank.org](http://www.emdatabank.org)) under accession numbers FMD-6108 to FMD-6117.

www.sciencemag.org/content/346/6211/857/suppl/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S6  
Tables S1 to S3  
References (25–59)

20 August 2014; accepted 14 October 2014  
10.1126/science.1260214



## VIRAL INFECTION

# Prevention and cure of rotavirus infection via TLR5/NLRC4-mediated production of IL-22 and IL-18

Benyue Zhang,<sup>1</sup> Benoit Chassaing,<sup>1</sup> Zhenda Shi,<sup>1</sup> Robin Uchiyama,<sup>1,2</sup> Zhan Zhang,<sup>1</sup> Timothy L. Denning,<sup>1,2</sup> Sue E. Crawford,<sup>3</sup> Andrea J. Pruijssers,<sup>4</sup> Jason A. Iskarpatyoti,<sup>4</sup> Mary K. Estes,<sup>3</sup> Terence S. Dermody,<sup>4,5</sup> Wenjun Ouyang,<sup>6</sup> Ifor R. Williams,<sup>2</sup> Matam Vijay-Kumar,<sup>7</sup> Andrew T. Gewirtz<sup>1,2\*</sup>

Activators of innate immunity may have the potential to combat a broad range of infectious agents. We report that treatment with bacterial flagellin prevented rotavirus (RV) infection in mice and cured chronically RV-infected mice. Protection was independent of adaptive immunity and interferon (IFN, type I and II) and required flagellin receptors Toll-like receptor 5 (TLR5) and NOD-like receptor C4 (NLRC4). Flagellin-induced activation of TLR5 on dendritic cells elicited production of the cytokine interleukin-22 (IL-22), which induced a protective gene expression program in intestinal epithelial cells. Flagellin also induced NLRC4-dependent production of IL-18 and immediate elimination of RV-infected cells. Administration of IL-22 and IL-18 to mice fully recapitulated the capacity of flagellin to prevent or eliminate RV infection and thus holds promise as a broad-spectrum antiviral agent.

Rotavirus (RV) causes severe dehydrating diarrhea in young children and moderate intestinal distress in adults (1). RV infection of adult mice serves as a well-defined model of intestinal viral infection in which infectivity can be monitored by measuring levels of viral antigen shed in feces (2). RV predominantly infects and replicates in epithelial cells lining the small intestine (1). Bacterial flagellin, the primary component of bacterial flagella, potentially activates host defense gene expression in intestinal epithelial cells (IECs) (3) and is viewed as a dominant immune activator in the intestine (4). Flagellin-induced gene expression confers both IECs and mice with resistance to a variety of challenges, including bacteria, chemicals, and radiation (5, 6). Moreover, administration of flagellin reduced susceptibility of mice to a culture-adapted strain of RV (7). Thus, we investigated the potential of flagellin to protect mice against highly contagious pathogenic mouse RV.

Oral inoculation of adult C57BL/6 mice with mouse RV (EC strain), using a dose 100,000 times that required to infect 50% of mice, resulted in uniform RV shedding that peaked 3 to 4 days

after inoculation and resolved several days later, indicating viral clearance (2). Administration of a single dose of flagellin via intraperitoneal injection shortly before infection temporarily prevented infection (fig. S1A). Repeated administration of flagellin further delayed (fig. S1B) or completely prevented infection (Fig. 1A). In contrast, repeated administration of bacterial lipopolysaccharide (LPS) only transiently protected against RV infection (fig. S2).

RV infection induces an adaptive immune response, which is normally required to clear infection (8). However, we saw very little elevation in antibodies to RV in flagellin-treated mice (fig. S3), suggesting that flagellin's protection against RV was independent of adaptive immunity. Confirming this notion, the protective effect of flagellin was fully maintained in mice lacking recombination activating gene 1 (*Rag1*<sup>-/-</sup>), making them deficient in mature B and T lymphocytes and susceptible to chronic RV infection (Fig. 1B) (9).

Analogous to young children, infection of neonatal mice with RV causes watery diarrhea (10). The incidence, duration, and active days of such RV-induced diarrhea and associated viral loads were reduced via flagellin treatment (Fig. 1, C and D). Flagellin treatment also diminished reovirus load after oral inoculation, suggesting broad antiviral activity (fig. S4).

We next investigated whether flagellin could treat established chronic RV infection in immune-compromised mice. Although *Rag1*<sup>-/-</sup> mice of all ages develop chronic RV infection, high viral loads are attained by infecting 3-week-old *Rag1*<sup>-/-</sup> mice (9). Flagellin treatment eliminated detectable RV shedding by 2 days after treatment. Moreover, 10 doses of flagellin treatment over a 20-day period abolished RV shedding for at least

150 days, whereas untreated mice shed virus over their lifetime (Fig. 1E). Accordingly, RV antigens, which were detectable by means of immunostaining in villus epithelial cells in untreated *Rag1*<sup>-/-</sup> mice, were absent in mice treated with flagellin by 24 hours (fig. S5A). Moreover, treatment with flagellin resulted in reduced levels of RV RNA within hours of administration, and RV RNA was undetectable 48 hours after treatment. Reduced levels of RV replication preceded a reduction in levels of RV genomic RNA (fig. S5B). In *Rag1*<sup>-/-</sup> mice, a substantial degree of replication occurs extra-intestinally, especially in the liver, resulting in hepatitis (11). Levels of RV RNA in the liver were greatly reduced within 24 hours of flagellin treatment and undetectable within 48 hours (fig. S5C). Similarly, flagellin treatment also eliminated the more modest levels of RV RNA in the spleen (fig. S5D). These results indicate that flagellin treatment cures chronic RV infection in immune-compromised mice.

Flagellin-induced remodeling of intestinal gene expression requires Toll-like receptor 5 (TLR5) (12). Flagellin can also be recognized intracellularly by the NOD-like receptor C4 (NLRC4) inflammasome, resulting in caspase 1-mediated production of the cytokines interleukin-1 $\beta$  (IL-1 $\beta$ ) and IL-18 (13, 14). We expected that only TLR5 would be required for flagellin protection against RV infection. However, loss of either pathway of flagellin recognition reduced, but did not completely eliminate, the capacity of flagellin to prevent RV infection (Fig. 2, A and B). In contrast, absence of both TLR5 and NLRC4 or myeloid differentiation primary response gene 88 (MyD88), which is required for signaling by TLR5 and inflammasome-associated cytokines, eliminated flagellin's protection against RV infection (Fig. 2, C and D).

Next, we generated bone marrow chimeric mice using wild-type (WT) and *Tlr5/Nlr4*<sup>-/-</sup> mice to determine the extent to which flagellin-mediated protection against RV infection requires recognition of flagellin by hemopoietic or nonhemopoietic cells. In contrast to our expectation that IEC recognition of flagellin would be essential, transplant of WT bone marrow to irradiated *Tlr5/Nlr4*<sup>-/-</sup> mice completely restored the capacity of flagellin to prevent RV infection, whereas administration of *Tlr5/Nlr4*<sup>-/-</sup> bone marrow to WT mice eliminated flagellin's anti-RV activity (fig. S6, A and B). Subsequent studies demonstrated that TLR5-dependent protection was restricted to hemopoietic cells, whereas NLRC4 expression by either cell compartment was sufficient for flagellin-mediated protection (fig. S6, C to F). This is consistent with reports of NLRC4 expression in macrophages and IECs (15, 16). Furthermore, hemopoietic and nonhemopoietic cells produce IL-18 in response to flagellin (fig. S7).

We next examined the role of hemopoietic cell types in mediating flagellin's antiviral effect. Although flagellin induced neutrophil recruitment to the intestine (fig. S8), neutrophils were not required for flagellin's antiviral effect (fig. S9). Nor were natural killer (NK) cells or intestinal macrophages required (figs. S10 and S11). Dendritic cells (DCs) mediate intestinal production of

<sup>1</sup>Center for Inflammation, Immunity and Infection, Institute for Biomedical Sciences, Georgia State University, Atlanta, GA, USA. <sup>2</sup>Department of Pathology, Emory University School of Medicine, Atlanta, GA, USA. <sup>3</sup>Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, TX, USA. <sup>4</sup>Elizabeth B. Lamb Center for Pediatric Research, Vanderbilt University School of Medicine, Nashville, TN, USA. <sup>5</sup>Departments of Pediatrics, Pathology, Microbiology, and Immunology, Vanderbilt University School of Medicine, Nashville, TN, USA. <sup>6</sup>Department of Immunology, Genentech, South San Francisco, CA, USA. <sup>7</sup>Department of Nutritional Sciences and Medicine, Pennsylvania State University, University Park, PA 16802, USA.

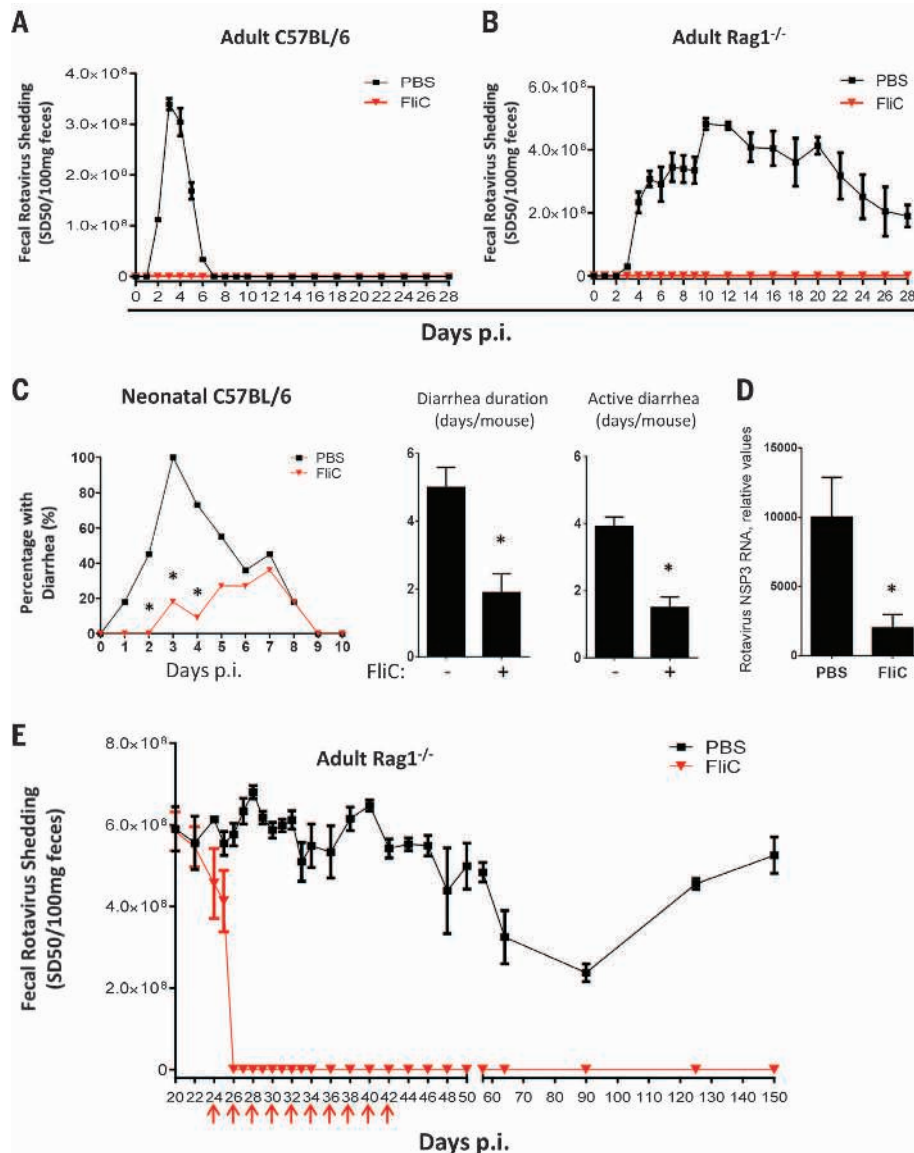
\*Corresponding Author. E-mail: agewirtz@gsu.edu



antibacterial peptides in response to flagellin (17). To investigate whether these cells play a role in flagellin's antiviral effect, we generated chimeric mice in which bone marrow-derived cells were engineered to express the diphtheria toxin receptor (DTR) under control of the CD11c promoter so that CD11c-expressing cells, primarily DCs, could be depleted by the administration of diphtheria toxin (DT) (18). Both DT-treated WT

mice and untreated chimeric mice were protected against RV infection by flagellin treatment (Fig. 2E and fig. S12). However, flagellin was unable to protect DT-treated chimeric mice against RV infection (Fig. 2E). Next, we investigated whether DCs expressing TLR5 and NLR4 are sufficient for flagellin protection against RV infection. *Tlr5/Nlr4*<sup>-/-</sup> mice were intravenously administered purified DCs [CD19<sup>-</sup>/major histocompat-

ibility complex (MHC) class II<sup>+</sup>/CD11c<sup>+</sup>/F4/80<sup>+</sup>] isolated from WT or *Tlr5/Nlr4*<sup>-/-</sup> mice. Only the WT DCs were capable of partial restoration of flagellin-mediated protection against RV infection, which was dependent on the number of DCs transferred (Fig. 2, F to H). Together, these results indicate that activation of TLR5/NLR4 on DCs is necessary and sufficient for flagellin's antiviral effect.

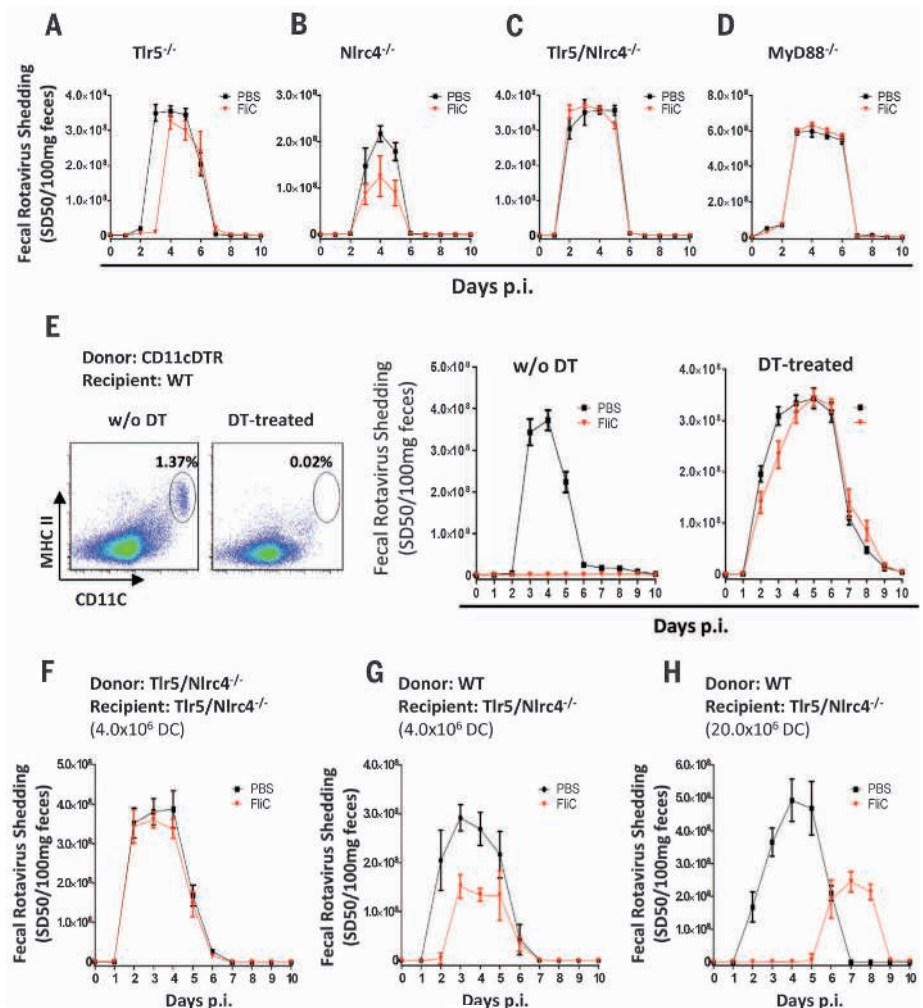


**Fig. 1. Flagellin protects mice from RV infection and diarrhea.** (A and B) Eight-week-old female (A) C57BL/6 mice or (B) *Rag1*<sup>-/-</sup> mice were orally inoculated with mouse RV, EC strain. Mice were administered 0.2 ml of phosphate-buffered saline (PBS) (vehicle) ± 20 µg of flagellin by means of intraperitoneal injection, and then every other day from 0 to 18 days after inoculation. Feces were collected daily and assayed for RV antigens by means of enzyme-linked immunosorbent assay (ELISA). Results are shown as mean ± SEM [two-way analysis of variance (ANOVA),  $n = 4$  mice,  $P < 0.001$ ] for (A) and (B). p.i., post-inoculation. (C) Seven-day-old C57BL/6 mice were orally inoculated with RV (supplementary materials). Mice were treated with PBS or flagellin (10 µg) every day from 0 to 9 days after inoculation and monitored for incidence of diarrhea daily ( $\chi^2$  test,  $n = 11$  mice,  $*P < 0.01$ ). Flagellin-treated

mice exhibited significantly reduced duration of diarrhea and days of active diarrhea (Student's  $t$  test,  $n = 11$  mice,  $*P < 0.01$ ). (D) Seven-day-old mice, treated as in (C), were euthanized 3 days after inoculation. Total RNA of small intestines from those mice were prepared and analyzed for RV NSP3 RNA level (Student's  $t$  test,  $n = 6$  mice,  $*P < 0.05$ ). (E) Three-week-old *Rag1*<sup>-/-</sup> mice were inoculated with mouse RV. Three weeks after inoculation, at which point a chronic infection had been established, mice were treated with PBS ± flagellin (20 µg) every 2nd day from 22 to 40 days after inoculation as indicated with red arrows. Feces were collected at indicated days and assayed for RV antigens by means of ELISA. Asterisk indicates that the difference between mice given PBS and flagellin was statistically significant. Results are shown as mean ± SEM (two-way ANOVA,  $n = 5$  mice,  $P < 0.0001$ ).

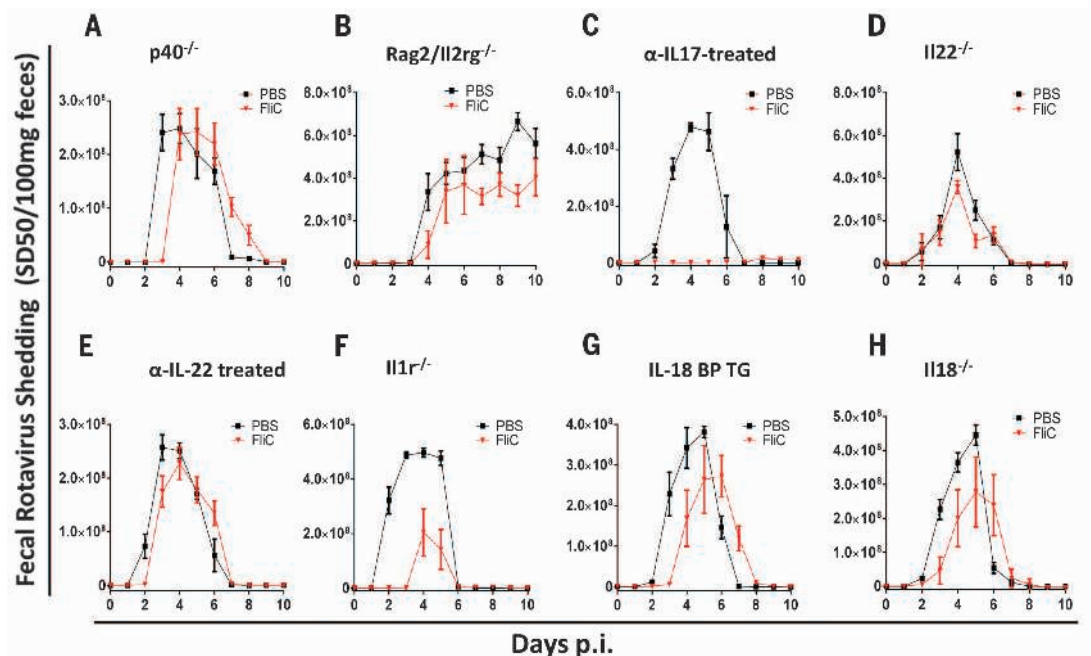


**Fig. 2. Flagellin's antiviral activity is mediated by TLR5/NLRC4 on dendritic cells.** (A to D) Eight-week-old (A) *Tlr5*<sup>-/-</sup>, (B) *Nlr4*<sup>-/-</sup>, (C) *Tlr5/Nlr4*<sup>-/-</sup>, and (D) *MyD88*<sup>-/-</sup> mice were orally inoculated with mouse RV and administered PBS ± flagellin (20 μg) every other day from 0 to 8 days after inoculation. Feces were assayed for RV antigens by means of ELISA. Results in (A) to (D) are shown as mean ± SEM [Student's *t* test, *n* = 5 mice, *P* < 0.05 on day 3 in (A) and day 5 in (B)]. Differences between PBS and flagellin groups in (C) and (D) were not significant [two-way ANOVA, *n* = 4 to 5 mice, *P* = 0.6361 for (C) and *P* = 0.3871 for (D)]. (E) CD11c-DTR reconstituted bone marrow chimeras were either untreated or injected with DT at 8 ng/gram body weight once a day for 2 days. Flow cytometry plots show the extent of depletion of DCs (CD45<sup>+</sup>/CD19<sup>-</sup>/MHC class II<sup>+</sup>/CD11c<sup>+</sup> splenocytes). Mice were then studied as in (A) to (D). The difference between the PBS and flagellin groups was statistically significant in absence of DT (two-way ANOVA, *n* = 4 mice, *P* < 0.001) and nonsignificant in DT-treated group (two-way ANOVA, *n* = 6 to 7 mice, *P* = 0.3821). (F and G) *Tlr5/Nlr4*<sup>-/-</sup> mice were adoptively transferred with 4 × 10<sup>6</sup> fluorescence-activated cell-sorted DCs (purity > 98.5%) from *Tlr5/Nlr4*<sup>-/-</sup> (F) or WT C57BL/6 (G) mice. Twelve hours later, the mice were studied as in (A) to (D). The difference between the PBS and flagellin groups was statistically significant in (G) (Student's *t* test on days 2 to 4, *n* = 4 mice, *P* < 0.05) but not (F) (Student's *t* test, *n* = 4 mice, *P* > 0.05 on all days). (H) *Tlr5/Nlr4*<sup>-/-</sup> mice were adoptively transferred with 20 million magnetic bead-sorted DCs (purity > 95.0%) from WT C57BL/6 mice, and 12 hours later, the mice were studied as in (A) to (D) (Student's *t* test on days 2 to 5, *n* = 4 mice, *P* < 0.05).



**Fig. 3. Flagellin-mediated protection against RV infection requires both IL-22 and IL-18.**

(A to H) Indicated strains of genetically modified 8-week-old mice were orally inoculated with mouse RV, EC strain. Mice were treated with PBS ± flagellin (20 μg), via intraperitoneal injection, every other day from 0 to 8 days after inoculation. Feces were collected daily and assayed for RV antigens by means of ELISA. The following strain was used in each panel: (A) *p40*<sup>-/-</sup>, (B) *Rag2/Il2rg*<sup>-/-</sup>, (C) WT C57BL/6 mice treated with IL-17-neutralizing mAb, (D) *Il22*<sup>-/-</sup>, (E) WT C57BL/6 mice treated with IL-22-neutralizing mAb, (F) *Il1r*<sup>-/-</sup>, (G) IL-18 BP TG, and (H) *Il18*<sup>-/-</sup> mice. Results in (A) to (H) are shown as mean ± SEM (*n* = 4 to 6 mice). The difference between mice given PBS and flagellin was statistically significant for (C) and (F) (two-way ANOVA, *P* < 0.0001) and significant at individual days of (A), (B), (D), (E), (G), and (H) [Student's *t* test, *P* < 0.05 on day 3 in (A), days 7 and 9 of (B), day 5 of (D), day 2 of (E), days 3 and 4 in (G), and days 3 and 4 in (H)].



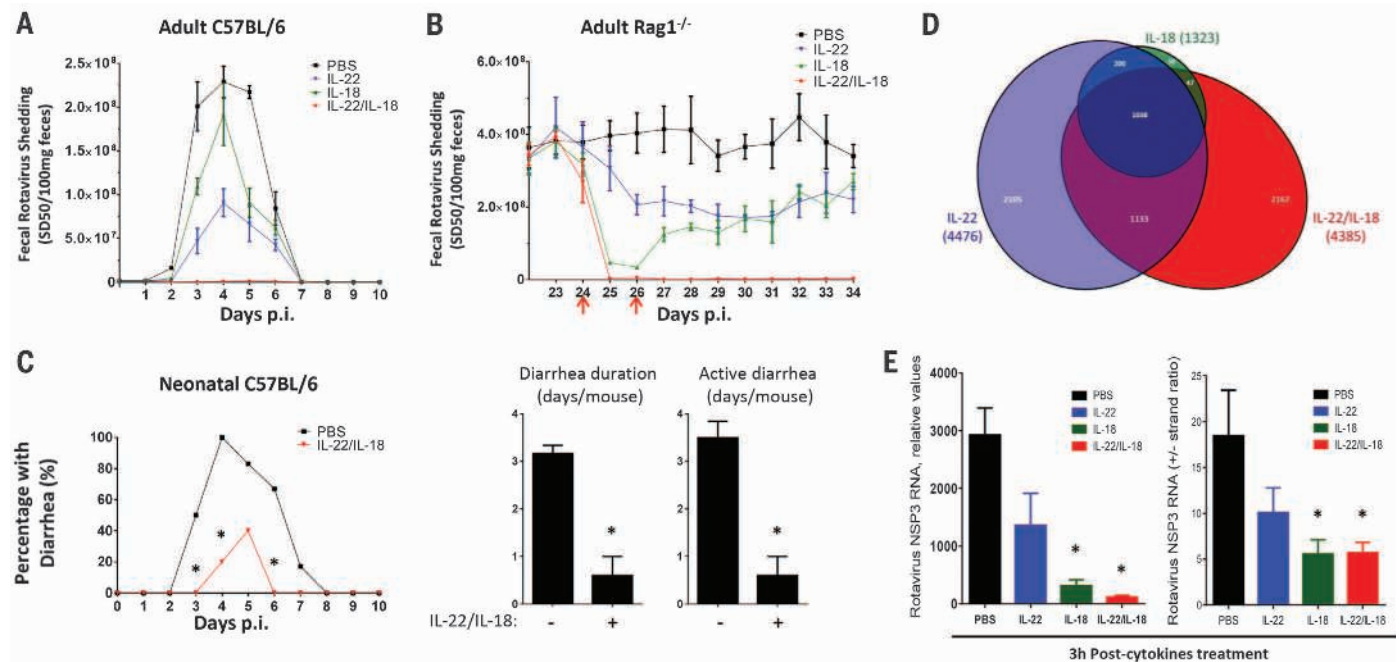
Interferons (IFNs) play a central role in antiviral immunity. However, flagellin's antiviral effect was fully maintained in mice lacking type I and/or II IFN receptors or in mice lacking signal transducer and activator of transcription 1 (STAT1), which mediates IFN signaling (fig. S13, A to D). Nor did flagellin induce expression of type III IFN (fig. S13E), which is induced by RV (19), or IFN-associated gene expression (20). Thus, flagellin prevents and clears RV infection by a previously unrecognized antiviral pathway.

DC TLR5 activates the IL-12/IL-23 axis, resulting in innate lymphoid cells' (ILCs) production of IL-17 and IL-22 (17, 21). Hence, we investigated whether ILC and this axis was involved in flagellin's antiviral action. Mice deficient in p40, which is a component of both IL-12 and IL-23, and *Rag2/Il2rg*<sup>-/-</sup> mice, which lack ILCs (and mature B and T lymphocytes), were not effectively protected by flagellin treatment (Fig. 3, A and B). Flagellin protection against RV infection was not affected by neutralization of IL-17 (Fig.

3C) but was almost completely abolished by genetic or antibody-mediated blockade of IL-22 (Fig. 3, D and E). These results suggest a central role for ILC-mediated IL-22 production in mediating flagellin's antiviral action. Accordingly, LPS does not elicit robust IL-22 production (27). Conversely, the requirement of NLR4 suggested possible roles for inflammasome cytokines IL-1 $\beta$  and IL-18. Ablation of IL-1 receptor signaling modestly impaired flagellin's protection against RV infection, whereas two different means of IL-18 ablation markedly reduced flagellin's antiviral effect (Fig. 3, F to H). In contrast to flagellin-induced IL-22 production, which requires ILCs, and DC expression of TLR5 (fig. S7) (22), flagellin-induced IL-18 production was unimpaired in *Rag2/Il2rg*<sup>-/-</sup> or DC-ablated mice (fig. S7), indicating that distinct signaling events drive IL-22 and IL-18 production in distinct cell types.

We next investigated the extent to which recombinant IL-22 and IL-18 recapitulate flagel-

lin's antiviral action. When administered prophylactically, IL-18 or IL-22 had only partial protective efficacy, but together conferred complete protection against a broad range of RV inoculation (Fig. 4A and fig. S14). Moreover, co-administration of IL-18 and IL-22 eliminated viral shedding in chronically RV-infected *Rag1*<sup>-/-</sup> mice (Fig. 4B). These results did not reflect a strong dependence on one of these cytokines for driving the expression of the other in that flagellin-induced IL-18 expression is independent of TLR5 (14), whereas flagellin-induced IL-22 is largely independent of NLR4 (fig. S15). Rather, our results suggest that parallel signaling pathways activated by IL-22 and IL-18 protect against RV infection and promote clearance of this virus, respectively. Consequently, combined treatment with IL-22 and IL-18 recapitulated the capacity of flagellin to cure RV infection in mice lacking mature T and B lymphocytes. Such combined IL-22/IL-18 treatment eliminated RV from *Rag1*<sup>-/-</sup> mice within 24 hours (versus 48 hours for flagellin) and, in



**Fig. 4. IL-22/IL-18 treatment prevents and treats RV infection.** (A) Eight-week-old C57BL/6 mice were treated with 0.2 ml PBS (vehicle), 2  $\mu$ g IL-22, 1  $\mu$ g IL-18, or 2  $\mu$ g IL-22 plus 1  $\mu$ g IL-18 by means of intraperitoneal injection and subsequently inoculated with RV. Such cytokines were administered every other day from 0 to 8 days after inoculation. Feces were assayed for RV antigens by means of ELISA. Results are shown as mean  $\pm$  SEM. Difference between PBS and IL-22/IL-18 groups were statistically significant (two-way ANOVA,  $n = 4$  mice,  $P < 0.0001$ ). (B) *Rag1*<sup>-/-</sup> mice, chronically infected with RV, were treated with PBS, 10  $\mu$ g IL-22, 1  $\mu$ g IL-18, or both on days 24 and 26 after inoculation (indicated with red arrows). Difference between PBS and IL-22/IL-18 groups was statistically significant (two-way ANOVA,  $n = 4$  mice,  $P < 0.0001$ ). (C) Seven-day-old C57BL/6 mice were orally inoculated with RV. Mice were administered 50  $\mu$ l PBS (vehicle) or 1  $\mu$ g IL-22 plus 0.2  $\mu$ g IL-18 immediately before inoculation, and 1 to 9 days after inoculation, and monitored for incidence of diarrhea daily ( $\chi^2$  test,  $n = 5$  to 6 mice,  $*P < 0.05$ ), duration and active days of diarrhea (Student's  $t$  test,  $n = 5$  to 6 mice,  $*P < 0.001$ ). (D to F) Chronically RV-infected *Rag1*<sup>-/-</sup> mice were treated with one injection of PBS, PBS containing 10  $\mu$ g IL-22, 1  $\mu$ g IL-18, or 10  $\mu$ g IL-22 plus 1  $\mu$ g IL-18. (D) Venn diagram representation of significant changes in intestinal epithelial gene expression 3 hours after cytokine treatment. (E) Intestinal levels of RV genomes and replication rates as reflected by NSP3 RNA levels and +/- RV strand ratios at 3 hours (Student's  $t$  test,  $n = 4$  mice,  $*P < 0.001$  for RV genome,  $*P < 0.05$  for RV RNA +/- strand ratio). (F) RV genomes levels at indicated time (Student's  $t$  test,  $n = 4$  mice,  $*P < 0.0001$ ).

contrast to flagellin, was effective in mice lacking ILCs (fig. S16). Like flagellin, prophylactic IL-22/IL-18 administration also afforded protection against the severe diarrhea seen in neonates (Fig. 4C). Furthermore, treating neonate mice with IL-22/IL-18 after diarrhea manifested moderately shortened its course (fig. S17). Thus, recapitulating flagellin's antiviral effect with IL-22/IL-18 might offer broad antiviral therapeutic potential even in immune-compromised hosts.

Neither flagellin nor IL-22/IL-18 altered RV infection in cultured IECs (fig. S18). Hence, to investigate mechanisms by which IL-22 and IL-18 treatment cleared RV infection, we examined signaling events in gut epithelial cells isolated from chronically RV-infected mice treated with IL-22, IL-18, or both cytokines. RNA sequencing revealed that IL-22 treatment induced rapid reprogramming of epithelial cell gene expression, implicating genes involved in a broad array of cellular processes (Fig. 4D and figs. S19 and S20). A more modest effect was observed in response to IL-18, whereas the combination of IL-22 and IL-18 induced a number of changes in gene expression not seen with either cytokine alone. In contrast, administration of IL-18 but not IL-22 to RV-infected mice resulted in rapid activation of caspase 3, suggesting involvement of apoptosis (fig. S21). Thus, IL-22 and IL-18 induced changes in gene expression and caspase 3 activation that correlated with rapid blockade of RV replication and elimination of RV genomes within 24 hours of cytokine treatment (Fig. 4, E and F). Together, these data suggest that IL-18 induces signaling events that lead to rapid reduction in RV levels, whereas IL-22 reprograms epithelial cell gene expression, resulting in resistance to RV infection.

We demonstrate that innate immunity can be harnessed to prevent and treat viral infection. Given the health burden caused by RV, including 600,000 annual deaths in children (23) and chronic infections in immune-compromised hosts, this strategy presents a therapeutic opportunity, differences between human and mouse RV strains (24) notwithstanding. Moreover, the complex antiviral action of flagellin was fully recapitulated by IL-22 and IL-18, whose use could circumvent differences in NLRC4 function between mice and humans (25). The action of these cytokines would likely be synergistic with therapies that directly target viruses and/or those that promote adaptive immunity. Thus, we propose activation of innate immunity with flagellin, or IL-22 and IL-18, as a potential strategy to combat emerging and recalcitrant viral pathogens.

## REFERENCES AND NOTES

- H. B. Greenberg, M. K. Estes, *Gastroenterology* **136**, 1939–1951 (2009).
- N. Feng, M. A. Franco, H. B. Greenberg, *Adv. Exp. Med. Biol.* **412**, 233–240 (1997).
- H. Zeng *et al.*, *J. Immunol.* **171**, 3668–3674 (2003).
- M. Vijay-Kumar, A. T. Gewirtz, *Mucosal Immunol.* **2**, 197–205 (2009).
- R. M. Jones *et al.*, *Gut* **60**, 648–657 (2011).
- L. G. Burdelya *et al.*, *Science* **320**, 226–230 (2008).
- M. Vijay-Kumar *et al.*, *J. Immunol.* **180**, 8280–8285 (2008).
- A. Sen, N. Feng, K. Ettayebi, M. E. Hardy, H. B. Greenberg, *J. Virol.* **83**, 10322–10335 (2009).
- J. Q. Jiang, X. S. He, N. Feng, H. B. Greenberg, *J. Virol.* **82**, 6812–6819 (2008).
- J. M. Ball, P. Tian, C. Q. Zeng, A. P. Morris, M. K. Estes, *Science* **272**, 101–104 (1996).
- I. Uhnou *et al.*, *J. Virol.* **64**, 361–368 (1990).
- F. A. Carvalho *et al.*, *Mucosal Immunol.* **5**, 288–298 (2012).
- A. Kupz *et al.*, *Nat. Immunol.* **13**, 162–169 (2012).
- M. Vijay-Kumar, F. A. Carvalho, J. D. Aitken, N. H. Fikadara, A. T. Gewirtz, *Eur. J. Immunol.* **40**, 3528–3534 (2010).
- L. Franchi *et al.*, *Nat. Immunol.* **7**, 576–582 (2006).
- S. Nordlander, J. Pott, K. J. Maloy, *Mucosal Immunol.* **7**, 775–785 (2014).
- M. A. Kinnebrew *et al.*, *Immunity* **36**, 276–287 (2012).
- S. Jung *et al.*, *Immunity* **17**, 211–220 (2002).
- J. Pott *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7944–7949 (2011).
- M. Vijay-Kumar *et al.*, *J. Immunol.* **174**, 6322–6331 (2005).
- L. Van Maele *et al.*, *J. Immunol.* **185**, 1177–1185 (2010).
- B. Chassaing, R. E. Ley, A. T. Gewirtz, *Gastroenterology* (2014).
- U. D. Parashar, C. J. Gibson, J. S. Bresee, R. I. Glass, *Emerg. Infect. Dis.* **12**, 304–306 (2006).
- J. W. Graff, K. Ettayebi, M. E. Hardy, *PLoS Pathog.* **5**, e1000280 (2009).
- Y. Zhao *et al.*, *Nature* **477**, 596–600 (2011).

## ACKNOWLEDGMENTS

We thank H. Virgin, T. Nice, A.-U. Rasheed, R. Ahmed, G. Silvestri, B. Rouse, B. Pulendran, A. Lukacher, S.-M. Kang, and D. Kalman for helpful discussions. We thank D. Moore and K. Yu for technical support. This work was supported by NIH grants DK061417—and an accompanying award from the American Reinvestment and Recovery Act—AI107943, DK064730, DK56338, AI038296, and AI080656. B.C. is supported by a fellowship from the Crohn's and Colitis Foundation of America. T.S.D. is supported by Elizabeth B. Lamb Center for Pediatric Research. Recombinant IL-22, antibodies to IL-17 and IL-22, and *Il22<sup>-/-</sup>* mice were provided under a materials transfer agreement by Genentech (South San Francisco, CA). Patents for use of flagellin and IL-22/IL-18 to treat or prevent viral infection have been applied for under application nos. PCT/US2011/052301 (Emory University) and 61/888/439 (Georgia State University, provisional). RNA-sequencing data are deposited in Gene Expression Omnibus with accession no. GSE62479.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/346/6211/861/suppl/DC1](http://www.sciencemag.org/content/346/6211/861/suppl/DC1)  
Materials and Methods  
Figs. S1 to S22  
References (26–45)

4 June 2014; accepted 16 October 2014  
10.1126/science.1256999



This is the start of something big.



**ScienceAdvances** |  AAAS  
SIGNIFICANT RESEARCH, GLOBAL IMPACT

Introducing *Science Advances* – the new, online-only, open-access journal from *Science* and AAAS. Find out how you can be among the first authors published at [scienceadvances.org](https://scienceadvances.org).

# Fidelity amplified.

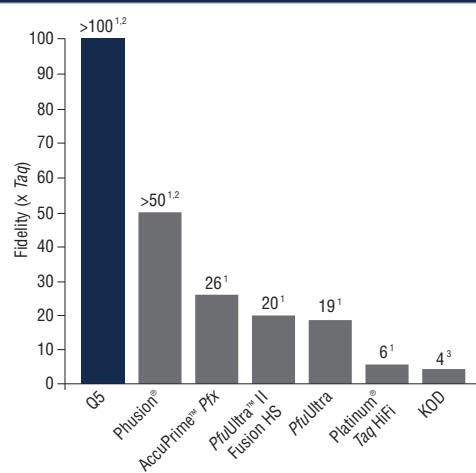
## Q5<sup>®</sup> High-Fidelity DNA Polymerase

Q5 High-Fidelity DNA Polymerase sets a new standard for both fidelity and performance. With the highest fidelity amplification available (>100X higher than *Taq*), PCR with Q5 DNA Polymerase results in ultra-low error rates. Its unique buffer system provides superior performance for a broad range of amplicons, regardless of GC content. Q5 DNA Polymerase represents the finest in fidelity.

**Request a free sample<sup>\*</sup>**  
**at [www.Q5PCR.com](http://www.Q5PCR.com)**

NEW ENGLAND BIOLABS<sup>®</sup>, NEB<sup>®</sup> and Q5<sup>®</sup> are registered trademarks of New England Biolabs, Inc. PHUSION<sup>®</sup> is a registered trademark and property of Thermo Fisher Scientific. Phusion<sup>®</sup> DNA Polymerase was developed by Finnzymes Oy, now a part of Thermo Fisher Scientific. PFULTRA<sup>™</sup> is a trademark of Agilent Technologies, Inc. PLATINUM<sup>®</sup> is a registered trademark of Life Technologies, Inc. ACCUPRIME<sup>™</sup> is a trademark of Life Technologies, Inc.

The highest fidelity amplification available



<sup>1</sup> PCR-based mutation screening in *lacZ* (NEB), *lacI* (Agilent) or *rpsL* (Life)

<sup>2</sup> Due to the very low frequency of misincorporation events being measured, the error rate of high-fidelity enzymes like Q5 is difficult to measure in a statistically significant manner. Although measurements from assays done side-by-side with Taq yield Q5 fidelity values from 100-200 X Taq, we report ">100X Taq" as a conservative value.

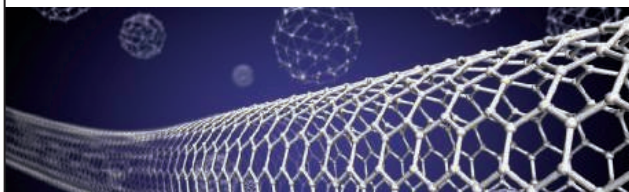
<sup>3</sup> Takagi et al (1997) *Appl. Env. Microbiol.* 63, 4504-4510.



## Biosystems Nanotechnology: Big Opportunities in the Science of the Small

The science of the very small is big business these days, as nanotechnology becomes a huge part of multiple sectors. In particular, scientists, engineers, and clinicians who endeavor to better understand how nanotechnology can impact biological systems—through the use of biosensors, biopharmaceuticals, and biomaterials—are finding abundant opportunities to pursue these investigations in multiple environments. Across the globe, demand is high in biosystems nanotechnology for professionals who speak the language of engineering and biology and have skill sets that include collaborating on diverse teams.

See the full story on page 870.



### Upcoming Features

*Regional Focus: Europe—November 28*

*Regional Focus: Asia—December 12*

*Faculty—January 30*

Produced by the Science/AAAS Custom Publishing Office

## AAAS Travels

### Alaska Aurora Borealis

March 26–April 1, 2015



**Come see the Greatest Light Show on Earth!**

See the dazzling night sky, spectacular snow-covered mountain peaks, grizzlies, moose, bald eagles, Denali (Mt. McKinley), and the famed Aurora Borealis (Northern Lights), the greatest light show on Earth! \$2,795 pp + air

**For a detailed brochure, call (800) 252-4910**

All prices are per person twin share + air



BETCHART EXPEDITIONS Inc.  
17050 Montebello Rd, Cupertino, CA 95014  
Email: AAASInfo@betchartexpeditions.com  
www.betchartexpeditions.com

## CALL FOR NOMINATIONS

# 2015

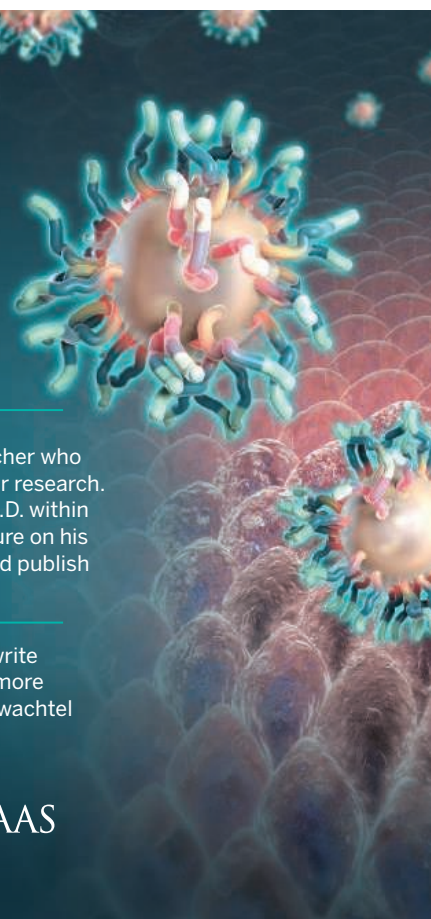
AAAS MARTIN AND  
ROSE WACHTEL  
CANCER RESEARCH

# AWARD

Recognize the work of an early-career cancer researcher who has performed outstanding work in the field of cancer research. Award nominees must have received their Ph.D. or M.D. within the last 10 years. The winner will deliver a public lecture on his or her research, receive a cash award of **\$25,000**, and publish a Focus article in *Science Translational Medicine*.

Submit nominations to [wachtelprize@aaas.org](mailto:wachtelprize@aaas.org) and write "Wachtel Award Nomination" in the subject line. For more information visit: [www.aaas.org/aboutaaas/awards/wachtel](http://www.aaas.org/aboutaaas/awards/wachtel)  
Deadline for submission: March 1, 2015.

Science Translational Medicine | AAAS







### Micromanipulator

The new TransferMan 4r is an innovative micromanipulator that combines an intuitive user interface with outstanding precision for rapid and simple microinjection. The exceptionally direct transmission of movement in all directions gives a "real-time feeling" and makes this instrument line the ideal platform for a broad range of applications, including microinjection and manipulation of suspension cells such as oocytes and blastocysts. The new DualSpeed joystick extends the precise and intuitive direct movement control with an additional dynamic speed mode for covering longer distances or speeding up sample processing. The multifunctional TransferMan 4r simplifies individual workflows by offering four predefined application-specific user interfaces such as "cell transfer" and "DNA injection." Users also have the option of programming additional methods for their individual application needs. The ergonomically designed control board and unique DualSpeed joystick mean the TransferMan 4r is tailored to providing excellent precision during cell handling and manipulation.

#### Eppendorf

For info: 800-645-3050

[www.eppendorf.com/cellmanipulation](http://www.eppendorf.com/cellmanipulation)

### LED Illumination System

The new Lumen 300-LED from Prior Scientific is the latest product in Prior's ever expanding microscopy illumination product line. The unit is specifically designed to offer broad spectrum LED illumination for fluorescence applications. Fitting directly to most microscopes, the system is simple to install and use. Controlled by a manual keypad controller, the Lumen 300-LED offers instant on/off operation via TTL, manual 0%–100% intensity control in 1% increments and on/off control for each LED. An optional liquid light guide is available when there is a need to keep the illumination source remote from the microscope. Providing 25,000+ hours of bulb life, the Lumen 300-LED is an environmentally friendly illumination system that requires no special disposal regulations.

#### Prior Scientific

For info: 800-877-2234

[www.prior.com](http://www.prior.com)

### Vacuum Jacketed Reactor System

The ReactoMate Reactor System, with glass vacuum jacket, offers outstanding thermal performance thereby improving control of reactions. By reducing heat loss/gain the glass vacuum jacket enables faster rates of heating and uses less energy to cool the system. Also, lower temperatures can potentially be achieved compared with vessels without a vacuum jacket. In addition, the vacuum jacket provides excellent reaction visibility and reduces condensation—staying ice-free to -50°C. The ReactoMate Reactor System is a safe and super-stable laboratory reactor. Compact in design, the Reacto-



### High-Capacity Slide Oven

The CytoBrite Slide Oven is a 60-slide convection incubator for performing extended or overnight FISH/ISH probe hybridizations. The CytoBrite Slide Oven has the capacity of five Thermobrite instruments in 60% less bench space. The oven has five pullout drawers that each receive two six-slide removable trays from the CytoBrite Slide Incubation System or CytoBrite Duo System. Slides are uniformly heated without humidification with the use of CytoBond Removable Coverslip Sealant.

#### SciGene

For info: 408-733-7337

[www.scigene.com](http://www.scigene.com)

Mate accommodates vacuum jacketed reactors from 50 mL to 10,000 mL and delivers precise control of reaction variables. Support systems for ReactoMate are constructed using stainless steel and aluminium components to ensure stability. Larger vessels (3,000–10,000 mL) are additionally supported at the base of the reactor to enhance safety. ReactoMate is designed to be compact, ergonomic, and have a simple setup.

#### Asynt

For info: +44-(0)-1638-781709

[www.asynt.com](http://www.asynt.com)

### Rotary Evaporator

The RC 900 Rotary Evaporator is designed to simplify rotary evaporation and enhance safety in the lab. The all-new RC 900 provides lab personnel with a unique combination of high-performance functionality and a simplified user experience, from an instrument that remains whisper-quiet and requires minimal space on the lab bench. The RC 900 provides lab personnel with streamlined rotary evaporation via a multitude of operational enhancements. Specifically, the instrument boasts straightforward flask exchange; now, a scientist can simply lock the flask into place using just one hand. Optimum flask angle is conveniently set via an adjustable knob. The system offers wireless control of all evaporation functions with an easy-to-use touchscreen and control knob. A cordless heating bath with pour spout encourages safe emptying without spilling, maximizing ease-of-use.

#### KNF Neuberger

For info: 609-890-8600

[www.knfusa.com](http://www.knfusa.com)

### Reaction Monitoring Probe

A new Direct Analysis Probe for the line of expression compact mass spectrometers (CMS) is now available. With hundreds of CMS systems utilized in chemistry laboratories around the world, this latest development creates yet another important tool for chemists that need fast, simple, and reliable sample analysis of liquids, solids, and tablets without the need for sample preparation. Recognizing the unrivaled flexibility of the Advion CMS as a detector across many workflows, the Direct Analysis Probe delivers essential mass spectral information of compounds in less than one minute. The new Direct Analysis Probe sample introduction, otherwise known as ASAP (Atmospheric Solids Analysis Probe), uses a standard glass melting point capillary which is introduced into the CMS system. Chemists dip the closed end of the capillary into the sample of interest or scrape the surface of a solid, and then place the capillary into the CMS for analysis.

#### Advion

For info: 877-523-8466

[www.advion.com](http://www.advion.com)

Electronically submit your new product description or product literature information! Go to [www.sciencemag.org/products/newproducts.dtl](http://www.sciencemag.org/products/newproducts.dtl) for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

# The 2015 Keystone Symposia Neurobiology Conferences

## Neuroinflammation in Diseases of the Central Nervous System

January 25–30, 2015

Sagebrush Inn and Conference Center | Taos, New Mexico | USA

Scientific Organizers: Richard M. Ransohoff, Christopher K. Glass and V. Hugh Perry

[www.keystonesymposia.org/15A5](http://www.keystonesymposia.org/15A5)

*Deadlines: Discounted Registration – Nov 24, 2014 (abstracts also still accepted online through this date)*

## Neuroepigenetics

February 22–26, 2015

Eldorado Hotel & Spa | Santa Fe, New Mexico | USA

Scientific Organizers: Hongjun Song and Li-Huei Tsai

[www.keystonesymposia.org/15B5](http://www.keystonesymposia.org/15B5)

*Deadlines: Abstract – Nov 19, 2014; Discounted Registration – Dec 18, 2014*

## Optogenetics

March 12–16, 2015

Westin Downtown Denver | Denver, Colorado | USA

Scientific Organizers: Edward S. Boyden, Klaus M. Hahn and Chandra Tucker

[www.keystonesymposia.org/15C5](http://www.keystonesymposia.org/15C5)

*Deadlines: Discounted Abstract/Scholarship – Nov 12, 2014; Abstract – Dec 11, 2014; Discounted Registration – Jan 13, 2015*

## Pathways of Neurodevelopmental Disorders

March 16–20, 2015

Granlibakken Resort | Tahoe City, California | USA

Scientific Organizers: Randi J. Hagerman, Mustafa Sahin and Paul J. Hagerman

[www.keystonesymposia.org/15C8](http://www.keystonesymposia.org/15C8)

*Deadlines: Discounted Abstract/Scholarship – Nov 18, 2014; Abstract – Dec 16, 2014; Discounted Registration – Jan 15, 2015*

## Neural Control of Metabolic Physiology and Diseases

April 12–17, 2015

Snowbird Resort | Snowbird, Utah | USA

Scientific Organizers: Dongsheng Cai and Martin G. Myers, Jr.

[www.keystonesymposia.org/15D4](http://www.keystonesymposia.org/15D4)

*Deadlines: Discounted Abstract/Scholarship – Dec 4, 2014; Abstract – Jan 13, 2015; Discounted Registration – Feb 12, 2015*

## Autophagy

June 19–24, 2015

Beaver Run Resort | Breckenridge, Colorado | USA

Scientific Organizers: Eric H. Baehrecke and Jayanta Debnath

[www.keystonesymposia.org/15E6](http://www.keystonesymposia.org/15E6)

*Deadlines: Discounted Abstract/Scholarship – Feb 17, 2015; Abstract – Mar 17, 2015; Discounted Registration – Apr 21, 2015*



Submitting an abstract represents an excellent way of participating in the conference via a poster presentation and possible selection for a short talk. Scholarships are available for graduate students and postdoctoral fellows. For full program, speaker, abstract and scholarship details, visit [www.keystonesymposia.org/neuro](http://www.keystonesymposia.org/neuro).

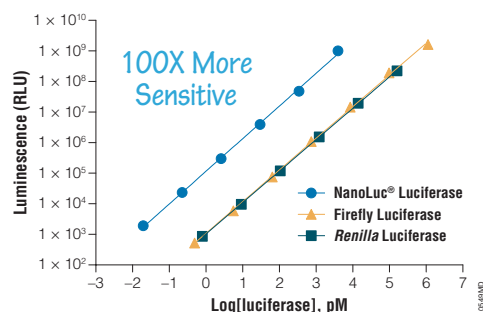
**KEYSTONE SYMPOSIA™**  
on Molecular and Cellular Biology  
*Accelerating Life Science Discovery*



# Redefine What Your Reporter Assays Can Do

## *Discover NanoLuc<sup>®</sup> Luciferase*

The small size and bright luminescence of NanoLuc<sup>®</sup> Luciferase bring exquisite sensitivity to traditional genetic reporter assays and expand your options for using a reporter to study biological interactions at physiologically relevant levels.



### *Discover how the brightness and sensitivity of NanoLuc<sup>®</sup> Luciferase can help you:*

- Detect gene regulation and cell signaling events
- Measure activity from endogenous promoters
- Detect protein:protein interactions in vivo
- Monitor protein stability
- Image events in cells or whole animals
- Identify lead compounds in target engagement studies

Get your **FREE NanoLuc<sup>®</sup> Vector:**  
**[www.promega.com/nano\\_sci](http://www.promega.com/nano_sci)**



**NanoDLR**  
Find out how  
the little things  
make all the  
difference.





### Micromanipulator

The new TransferMan 4r is an innovative micromanipulator that combines an intuitive user interface with outstanding precision for rapid and simple microinjection. The exceptionally direct transmission of movement in all directions gives a "real-time feeling" and makes this instrument line the ideal platform for a broad range of applications, including microinjection and manipulation of suspension cells such as oocytes and blastocysts. The new DualSpeed joystick extends the precise and intuitive direct movement control with an additional dynamic speed mode for covering longer distances or speeding up sample processing. The multifunctional TransferMan 4r simplifies individual workflows by offering four predefined application-specific user interfaces such as "cell transfer" and "DNA injection." Users also have the option of programming additional methods for their individual application needs. The ergonomically designed control board and unique DualSpeed joystick mean the TransferMan 4r is tailored to providing excellent precision during cell handling and manipulation.

#### Eppendorf

For info: 800-645-3050

[www.eppendorf.com/cellmanipulation](http://www.eppendorf.com/cellmanipulation)

### LED Illumination System

The new Lumen 300-LED from Prior Scientific is the latest product in Prior's ever expanding microscopy illumination product line. The unit is specifically designed to offer broad spectrum LED illumination for fluorescence applications. Fitting directly to most microscopes, the system is simple to install and use. Controlled by a manual keypad controller, the Lumen 300-LED offers instant on/off operation via TTL, manual 0%–100% intensity control in 1% increments and on/off control for each LED. An optional liquid light guide is available when there is a need to keep the illumination source remote from the microscope. Providing 25,000+ hours of bulb life, the Lumen 300-LED is an environmentally friendly illumination system that requires no special disposal regulations.

#### Prior Scientific

For info: 800-877-2234

[www.prior.com](http://www.prior.com)

### Vacuum Jacketed Reactor System

The ReactoMate Reactor System, with glass vacuum jacket, offers outstanding thermal performance thereby improving control of reactions. By reducing heat loss/gain the glass vacuum jacket enables faster rates of heating and uses less energy to cool the system. Also, lower temperatures can potentially be achieved compared with vessels without a vacuum jacket. In addition, the vacuum jacket provides excellent reaction visibility and reduces condensation—staying ice-free to  $-50^{\circ}\text{C}$ . The ReactoMate Reactor System is a safe and super-stable laboratory reactor. Compact in design, the Reacto-



### High-Capacity Slide Oven

The CytoBrite Slide Oven is a 60-slide convection incubator for performing extended or overnight FISH/ISH probe hybridizations. The CytoBrite Slide Oven has the capacity of five Thermobrite instruments in 60% less bench space. The oven has five pullout drawers that each receive two six-slide removable trays from the CytoBrite Slide Incubation System or CytoBrite Duo System. Slides are uniformly heated without humidification with the use of CytoBond Removable Coverslip Sealant.

#### SciGene

For info: 408-733-7337

[www.scigene.com](http://www.scigene.com)

Mate accommodates vacuum jacketed reactors from 50 mL to 10,000 mL and delivers precise control of reaction variables. Support systems for ReactoMate are constructed using stainless steel and aluminium components to ensure stability. Larger vessels (3,000–10,000 mL) are additionally supported at the base of the reactor to enhance safety. ReactoMate is designed to be compact, ergonomic, and have a simple setup.

#### Asynt

For info: +44-(0)-1638-781709

[www.asynt.com](http://www.asynt.com)

### Rotary Evaporator

The RC 900 Rotary Evaporator is designed to simplify rotary evaporation and enhance safety in the lab. The all-new RC 900 provides lab personnel with a unique combination of high-performance functionality and a simplified user experience, from an instrument that remains whisper-quiet and requires minimal space on the lab bench. The RC 900 provides lab personnel with streamlined rotary evaporation via a multitude of operational enhancements. Specifically, the instrument boasts straightforward flask exchange; now, a scientist can simply lock the flask into place using just one hand. Optimum flask angle is conveniently set via an adjustable knob. The system offers wireless control of all evaporation functions with an easy-to-use touchscreen and control knob. A cordless heating bath with pour spout encourages safe emptying without spilling, maximizing ease-of-use.

#### KNF Neuberger

For info: 609-890-8600

[www.knfusa.com](http://www.knfusa.com)

### Reaction Monitoring Probe

A new Direct Analysis Probe for the line of expression compact mass spectrometers (CMS) is now available. With hundreds of CMS systems utilized in chemistry laboratories around the world, this latest development creates yet another important tool for chemists that need fast, simple, and reliable sample analysis of liquids, solids, and tablets without the need for sample preparation. Recognizing the unrivaled flexibility of the Advion CMS as a detector across many workflows, the Direct Analysis Probe delivers essential mass spectral information of compounds in less than one minute. The new Direct Analysis Probe sample introduction, otherwise known as ASAP (Atmospheric Solids Analysis Probe), uses a standard glass melting point capillary which is introduced into the CMS system. Chemists dip the closed end of the capillary into the sample of interest or scrape the surface of a solid, and then place the capillary into the CMS for analysis.

#### Advion

For info: 877-523-8466

[www.advion.com](http://www.advion.com)

Electronically submit your new product description or product literature information! Go to [www.sciencemag.org/products/newproducts.dtl](http://www.sciencemag.org/products/newproducts.dtl) for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.



There's only one **Science**

## Science Careers Advertising

For full advertising details, go to  
ScienceCareers.org and click  
For Employers, or call one of  
our representatives.

### Tracy Holmes

Worldwide Associate Director  
Science Careers  
Phone: +44 (0) 1223 326525

### THE AMERICAS

E-mail: [advertise@sciencecareers.org](mailto:advertise@sciencecareers.org)  
Fax: 202 289 6742

### Tina Burks

Phone: 202 326 6577

### Nancy Toema

Phone: 202 326 6578

### Marci Gallun

Sales Administrator  
Phone: 202 326 6582

### Online Job Posting Questions

Phone: 202 312 6375

### EUROPE / INDIA / AUSTRALIA / NEW ZEALAND / REST OF WORLD

E-mail: [ads@science-int.co.uk](mailto:ads@science-int.co.uk)  
Fax: +44 (0) 1223 326532

### Axel Gesatzki

Phone: +44 (0) 1223 326529

### Sarah Lelarge

Phone: +44 (0) 1223 326527

### Kelly Grace

Phone: +44 (0) 1223 326528

### JAPAN

Katsuyoshi Fukamizu (Tokyo)

E-mail: [kfukamizu@aaas.org](mailto:kfukamizu@aaas.org)  
Phone: +81 3 3219 5777

Hiroyuki Mashiki (Kyoto)

E-mail: [hmashiki@aaas.org](mailto:hmashiki@aaas.org)  
Phone: +81 75 823 1109

### CHINA / KOREA / SINGAPORE / TAIWAN / THAILAND

### Ruolei Wu

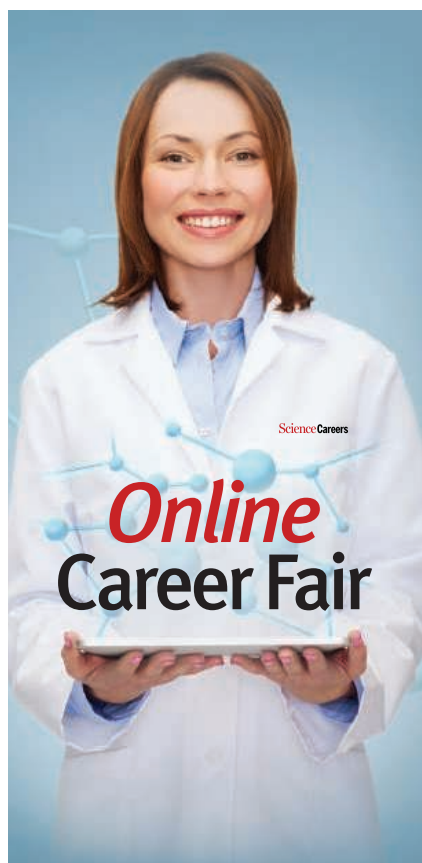
Phone: +86 186 0082 9345  
E-mail: [rwu@aaas.org](mailto:rwu@aaas.org)

All ads submitted for publication must comply with applicable U.S. and non-U.S. laws. Science reserves the right to refuse any advertisement at its sole discretion for any reason, including without limitation for offensive language or inappropriate content, and all advertising is subject to publisher approval. Science encourages our readers to alert us to any ads that they feel may be discriminatory or offensive.

**Science Careers**

FROM THE JOURNAL SCIENCE AAAS

[ScienceCareers.org](http://ScienceCareers.org)



## Online Career Fair

Meet recruiters  
without leaving  
your lab or desk!

December 10, 2014

Science Careers is excited to  
bring you an easy, efficient and  
powerful way to connect with  
employers from all over... all  
in one place.

For more information, visit:  
[ScienceCareers.org/  
onlinecareerfair](http://ScienceCareers.org/onlinecareerfair)

For careers in science,  
there's only one **Science**

**Science Careers**

FROM THE JOURNAL SCIENCE AAAS

For information on recruiting  
at this event, please visit:

[ScienceCareers.org/  
onlinecareerfairemployers](http://ScienceCareers.org/onlinecareerfairemployers)



## Tenure-Track Assistant Professorship in Chemistry and Chemical Biology Harvard University Cambridge, Massachusetts

Candidates are invited to apply for an open-field tenure-track assistant professorship in the Department of Chemistry and Chemical Biology at Harvard University. The appointment is expected to begin on July 1, 2015. The tenure-track professor will be responsible for teaching at the undergraduate and graduate levels. We are seeking candidates who have an outstanding research record and a strong commitment to undergraduate and graduate teaching. Doctorate required by expected start date. Candidates should arrange to have three letters of recommendation sent independently and provide a curriculum vitae, statement of teaching philosophy, list of publications, and outline of their future research plans. All applications and supporting materials must be submitted via the ARIES portal (<http://academicpositions.harvard.edu/postings/5829>) no later than **December 15, 2014**.

*Harvard is an Equal Opportunity Employer and all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability status, protected veteran status, or any other characteristic protected by law.*

## Science Careers Cernet

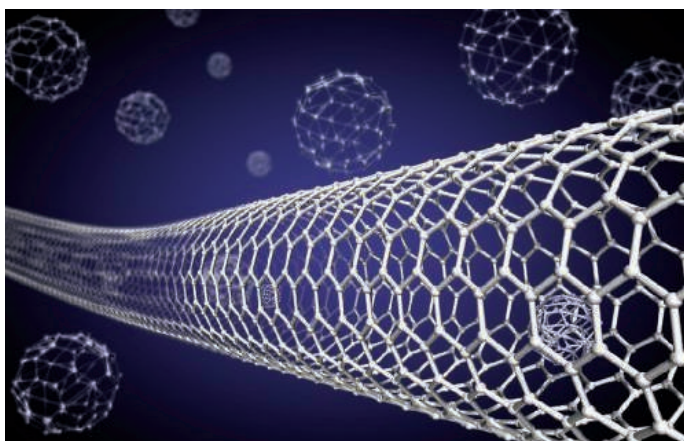
“《科学》职业” 已  
经与Cernet/赛尔互  
联开展合作。中国大  
陆的高校可以直接联  
系Cernet/赛尔互联  
进行国际人才招聘。



请访问  
[Sciencecareers.org/CER](http://Sciencecareers.org/CER)  
点得联系信息。

**Science**





## Biosystems Nanotechnology: Big Opportunities in the Science of the Small

The science of the very small is big business these days, as nanotechnology becomes a huge part of multiple sectors. In particular, scientists, engineers, and clinicians who endeavor to better understand how nanotechnology can impact biological systems—through the use of biosensors, biopharmaceuticals, and biomaterials—are finding abundant opportunities to pursue these investigations in multiple environments. Across the globe, demand is high in biosystems nanotechnology for professionals who speak the language of engineering and biology and have skill sets that include collaborating on diverse teams. **By Alaina G. Levine**

**Chris Skipwith** was looking for a way to make his life sciences and physics research more meaningful and move ideas more quickly into therapeutics. With a doctorate in biophysics and experience in pharmaceuticals as a result of a lengthy internship at Merck, he knew that serotonin levels change in patients with thrombosis, which can cause dangerous clots to develop in the bloodstream. Although anti-coagulant medications exist, many of them require regular blood sample collection for monitoring thrombotic risk, and Skipwith saw a problem that he might be able to help solve. He became interested in applying his background in imaging techniques and X-ray crystallography to develop a therapeutic solution that relied on nanotechnology—a real-time biosensor that could continuously and non-invasively monitor someone's blood serotonin, and help them fight off blood clots even before they form. But before he could attack

this problem he knew he needed assistance, as he wasn't a nanotech expert.

In exploring potential research collaborators, he came across **Heather Clark**, a professor of pharmaceutical sciences at Northeastern University's Bouvé College of Health Sciences, who was working on numerous biosensor-related projects at the nanotech scale. Skipwith contacted her with a request for a postdoc appointment and a collaboration was born, as Clark immediately realized the value they both could provide each other. "He taught us techniques in biophysics, and we taught him about nanoscience," she says. Today, as Clark's postdoc, he partners with chemists, pharmaceutical scientists, and engineers in her lab as he develops the biosensors.

This research "marriage" is not atypical in biosystems nanotechnology labs: Across academia, government, and industry, groups are almost always interdisciplinary, and new employees—whether they are postdocs or permanent staff—are hired based on how they can holistically contribute to the team, or for specific skills they possess which will complement the group's expertise.

Biosystems nanotechnology, or nanomedicine, requires diversity amongst its researchers because of its complexity. Medical devices, pharmaceuticals, and sensors can have nanotechnological elements or can be built at the nanoscale themselves. For example, "we're taking off-the-shelf [medical devices], such as hip implants, which typically have a failure rate after 20 years, and putting nanomaterials on the surface," explains **Tom Webster**, chair and professor of chemical engineering at Northeastern University and president-elect of the Society for Biomaterials. This nanofication can decrease inflammation and scar tissue and encourage bone growth. "It's not changing the chemistry of the implant, but incorporating nanotexturing onto the surface of the device itself." Scientists see enormous growth potential in adding nanoparticles to stents, orthopedics, catheters, and even dental implants to speed the healing process. "Some people think that nanomedicine is 20 years away, but short-term examples (like medical devices with nanofeatures) are happening now," he says.

### Global contributions

Indeed, "a huge number of innovations associated with problems that nanotechnology can solve are in the healthcare space," echoes **Lloyd Whitman**, interim director of the (U.S.) National Nanotechnology Initiative (NNI) coordination office, which manages the activities under the U.S. National Nanotechnology Initiative (NNI). He notes that the largest budget for nanotech projects in the U.S. government currently resides in the NIH, at \$441.5 million per year.

In fact, across the world, countries are recognizing the potential of this field to radically change health care. Asia especially is showing growth, as nations such as Singapore, South Korea, Taiwan, Japan, India, and China invest in research in these areas. "A lot of countries are focused on physical science applications to biosystems," says **Saion Sinha**, a professor of physics and electrical engineering at the University of New Haven in Connecticut and a researcher in nanomaterials. "There used to be silos where biologists and physicists didn't talk. But Asian governments are playing a role in building centers that [encourage collaboration]. **continued>**

### Upcoming Features

Regional Focus: Europe—November 28 ■ Regional Focus: Asia—December 12 ■ Faculty—January 30

# Postdoctoral Program

Novartis Institutes for  
BioMedical Research

Do you have a passion for innovative  
fundamental research in drug discovery?

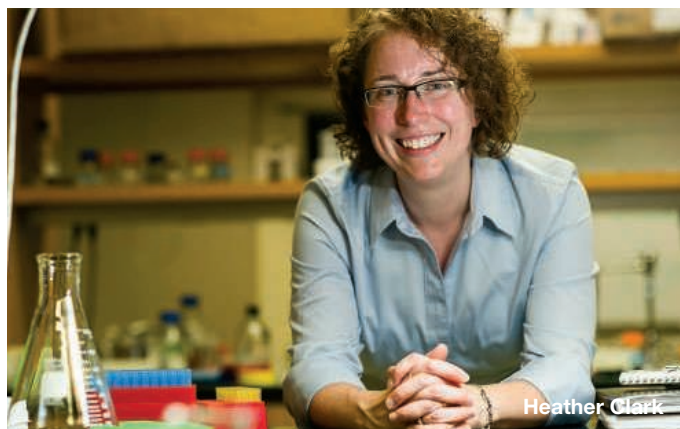
We are seeking creative postdoctoral scholars to  
conduct exciting research in biology, chemistry,  
and computational sciences at the frontier of  
drug discovery.

Postdocs at NIBR who will subsequently pursue  
faculty positions in academia are eligible for  
start-up funds through our new and highly  
competitive Young Investigator Awards.

Visit <http://postdoc.nibr.com> to view our  
mentors' research profiles and to apply.







**Proactively reach out to leaders and make yourself, your interests, and skills known, advises Clark.**

That's why Asia is forging ahead [in nanomedicine]."

Singapore has shown to be a powerhouse of nanotech research as "it is pouring lots of money into recruiting high-level scientists in nanomedicine and nanotechnology," says **Scott McNeil**, who heads the U.S. National Cancer Institute's Nanotechnology Characterization Laboratory (NCL). The nation began ramping up its investments into biomedical research around 2000, notes **Chwee Teck Lim**, an entrepreneur and provost's chair professor in the departments of biomedical engineering and mechanical engineering at the National University of Singapore (NUS). "It's an exciting time in Singapore because the government is making available lots of money to do research and it is also very pro-R&D," he says. Its National Research Foundation, only eight years old, provides hefty fellowships to young investigators (up to US\$3 million), as well as support for technology accelerators and startups. Furthermore, it offers financial incentives for overseas universities to set up shop in the country.

Taiwan is in a current state of enhancing its abilities to contribute to biosystems nanotechnology, says **Tao-Shih Hsieh**, a distinguished research fellow and director of the Institute of Cellular and Organismic Biology at Academia Sinica in Taipei. "Medical science is the next phase of Taiwanese technology [advancement]," he says. The government is investing in its major universities, and alumni and industry colleagues are actively being recruited back to the nation to participate in research activities associated with this field. "Taiwan is in a position to be much stronger in nanotechnology in the future," he concludes.

Similarly, Japan seems to be trying hard to expand its reach in biosystems nanotech. The country's 2006–2010 Science & Technology Plan identified nanotechnology and new materials as part of eight promotion areas and helped establish a new section of Japan's National Institute of Health Sciences (NIHS) that focuses on nanomedicine. Both the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the Ministry of Health, Labour and Welfare (MHLW) (which oversees the NIHS) have invested in projects in nanotechnol-

ogy in biological systems. Although experts agree that the country is investing in nanotechnology in general, there is still much to do in terms of solidifying Japan's reputation as a nanomedicine leader. **Toru Maekawa**, the director of the Bi-Nano Electronics Research Center at Toyo University, notes that his organization, which was established in 1996, has received MEXT funding, but the strongest government support is still in traditional research areas and not frontier science like biosystems nanotechnology. "I hope Japan will be very good at nanomedicine [in the future]," he says.

China is in a growth spurt. **Yuliang Zhao**, deputy director-general of the National Center for Nanosciences and Technology of China, and director of the Chinese Academy of Sciences (CAS) Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety, sees a number of actions taking place in the nation currently that will impact its prominence in the field of nanomedicine, including "the higher-and-higher-quality academic activities, the long-term governmental investment, the ambitious venture capital, the transformative industry (from labor-intensive to tech-intensive), and the vast needs from a huge population of patients," he says. "All of these are moving together and their integration will impact and promote the uses of nanotech in biosystems and medical technology in the future." That being said, he and other experts suggest that China still has a ways to go before it is considered to be a top-tier contributor to this arena on the world stage.

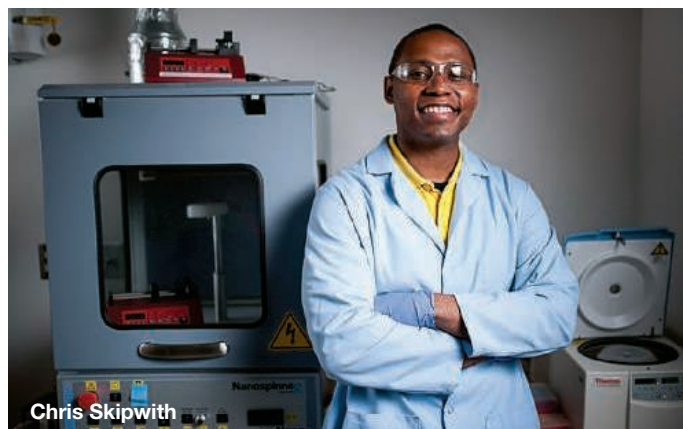
In the European Union, continental programs are pushing to make an impact in nanomedicine. The European Commission is in the process of creating an E.U.-wide nanotech characterization laboratory, similar to the NCL in the United States, says McNeil. The new center will serve small companies and academic labs by not only conducting characterization studies, but also pursuing opportunities for scaling up the technological innovations, the latter of which "is the missing piece in the U.S.," he says. The latest E.U. research funding framework, Horizon 2020, has provisions that support applied and industry-related investigations, which can include biosystems nanotech.

**Sonia Contera**, a Spanish physicist who co-directs the Oxford Martin School's Institute of Nanoscience for Medicine, sees Horizon 2020 as a potential game-changer in the discipline. "Individual country's governments are not funding individual scientists like they used to, and people like me, who are multidisciplinary, don't fit with local country funding programs," she says. "But Horizon 2020 makes it easier to obtain grants and will lead to more technological development," as it funds research that aligns scientists and engineers with clinicians working at hospitals and in industry.

### Group makeup and leadership

One key component of successful medical nanotechnology programs is the diverse makeup of the teams. The U.S. National Cancer Institute (NCI) Alliance for Nanotechnology in Cancer funds nine academic centers that wed nanotechnology and biological systems, and since its beginning, it has mandated that its centers are co-led by physical scientists or engineers and cancer biologists or oncologists. "This dual leadership works quite well because of the sophistication and diversity of expertise needed in the medical nanotech space," says its director, **Piotr Grodzinski**.

Beyond the NCI, many programs are collaboratively



Chris Skipwith

### Featured Participants

Academia Sinica  
www.sinica.edu.tw

Boston Scientific  
www.bostonscientific.com

CAS Key Laboratory  
for Biomedical Effects  
of Nanomaterials and  
Nanosafety  
nanosafety.cas.cn

Johnson & Johnson  
www.jnj.com

National Center for  
Nanosciences and  
Technology of China  
nanoctr.cas.cn

Nanotechnology  
Characterization Lab  
ncl.cancer.gov

National Nanotechnology  
Initiative  
www.nano.gov

National University of  
Singapore  
www.nus.edu.sg

NCI Alliance for  
Nanotechnology in Cancer  
nano.cancer.gov

Northeastern University  
www.northeastern.edu

Oxford Martin School,  
University of Oxford  
www.oxfordmartin.ox.ac.uk

Society for Biomaterials  
www.biomaterials.org

Toyo University  
www.toyo.ac.jp/site/english/  
index.html

University of California San  
Diego  
www.ucsd.edu

University of New Haven  
www.newhaven.edu

managed by life and physical scientists. Case in point: **Shao-chen Chen**, a professor of nanoengineering and bioengineering at the University of California San Diego, who co-directs the university's Biomaterials and Tissue Engineering Center (which is part of the Institute of Engineering in Medicine) with a medical doctor. "The problems require us to work together. As engineers, we have to find real biological needs for the tools we develop, otherwise they will be useless," he shares. "We don't know the critical biological questions, but the medical researchers don't know how to build the tools that we do. The engineers, biologists, and clinicians have to work together."

Lim's team sits in the NUS Mechanobiology Institute and the faculty of engineering, and as he consults with biologists and clinicians as dictated by the nature of his various projects, he stresses the need for engineers and physical scientists to partner with medical doctors early and regularly in the innovation

process. "Often engineers will come up with an idea that clinicians do not really need," i.e., they develop a solution looking for a problem, he says. "They need to interact with each other right at the very early stage—that's the bottom line. It could be the engineer leading the project or the clinician, but either way it's important to collaborate and get off on the right foot."

### Where the jobs are

Given that nanotechnology experts come from a plethora of fields, their job opportunities are just as varied and rich. In academia, departments such as chemical engineering, materials science, physics, and chemistry are hiring, as are pharmaceutical sciences, medicine, and various life sciences.

In industry, nanotech biosystems investigations may take place in R&D or manufacturing, says Webster. This is the structure found at Johnson & Johnson, says **Ibraheem Badejo**, a biomaterials scientist with J&J's Cambridge Innovation Center in Massachusetts, where nanotech experts are found in R&D or in the preclinical/product development side. Physicists and mechanical and chemical engineers tend to start in research, and biologists with nanotechnology experience often start on the preclinical side, he adds.

At Boston Scientific, nanotechnology experts find themselves essentially as consultants for projects throughout the organization, describes **Peter Edelman**, a biomaterials expert and R&D fellow. His projects range from exploratory research to support and expansion of existing products, but scientists and engineers with nanotechnology knowledge could find themselves in any part of the firm, including product development and manufacturing. As "big pharma is actively seeking out investigations to do reformulations of existing small molecules into nanoparticles," notes McNeil, there has been an uptick in career opportunities in both big and small firms in this sector, including startups.

### Advice for advancement

Webster predicts that as nanotech takes a larger role in medicine, there will be more opportunities available for those who have nanotechnology experience. "All of the exciting research taking place in biomaterials and biopharmaceuticals is at the nanoscale," he says. "There are [plenty] of opportunities for learning and research—it's a wonderful [time] to embrace the field."

So how do you prepare for these openings? "Learn the language and culture of the different fields so you can bridge them," says Whitman. For life scientists, opportunities will be significantly broader as they expand their horizons and hone skills in techniques and tactics employed by engineers and physical scientists. "When scientists are cross-trained in more than one discipline, they become more marketable," says Grodzinski.

Proactively reach out to leaders and make yourself, your interests, and skills known, advises Clark. And since nanotech is global in it of itself, experience working in other cultures and nations will bolster your resume, says Chen. But perhaps most importantly, keep nurturing your inquisitive nature. As Whitman encourages his protégés to ask themselves: "Where are the questions that haven't been answered?"

*Alaina G. Levine is a science and careers writer based in Tucson, AZ.*

DOI: 10.1126/science.opms.r1400149





THE UNIVERSITY OF TEXAS  
AT DALLAS  
The Erik Jonsson School of  
Engineering and Computer Science

## BIOENGINEERING

### ENDOWED FACULTY POSITIONS

The Erik Jonsson School of Engineering and Computer Science at the University of Texas at Dallas (UTD) invites applications for two endowed chair positions at the Full Professor level in Bioengineering and/or Biomedical Engineering. Candidates from related disciplines, such as Electrical Engineering, Mechanical Engineering, Computer Science, or Materials Science, whose research is relevant to Bio/Biomedical Engineering, may also be considered. Successful candidates are expected to be internationally renowned researchers and educators in the field with a strong commitment to building a world-class research program at UTD. Of particular interest are candidates whose research interests span one or more of the following areas: biomedical devices, biomaterials, biomechanics, neural engineering, systems biology, computational biology, imaging, medical robotics, including surgical robots, and neurosciences.

The Department of Bioengineering at UTD currently offers the BS, MS and Ph.D. degrees in Biomedical Engineering. The MS and Ph.D. degree programs are offered jointly with The University of Texas Southwestern Medical Center at Dallas and the University of Texas at Arlington. The close proximity to the Medical School offers outstanding opportunities for collaborative research. In addition, a \$108M Bioengineering and Sciences Building is currently under construction and will open on the UTD campus in Fall, 2015. The newly established Texas Biomedical Device Center at UTD offers further opportunities for translational research.

In addition to Bioengineering, the Erik Johnson School is home to the Departments of Electrical Engineering, Computer Science, Mechanical Engineering, Material Science & Engineering, and Systems Engineering, and has interdisciplinary programs in Computer Engineering, Telecommunications Engineering, and Software Engineering.

The University of Texas at Dallas is situated in Richardson, one of the most attractive suburbs of the Dallas metropolitan area with several hundred high-tech companies within a few miles of the campus, including Texas Instruments, Alcon, Abbot, Greatbatch, GE Healthcare, Advanced Arm Dynamics, Stryker, VLSIP, ZyveX, Samsung, BlueCross/BlueShield, Intervoice, Plexon, Raytheon, Lennox, Ericsson, Hewlett-Packard, Samsung, Fujitsu, Cisco Systems, EDS, and ZyveX. Opportunities for joint university-industry research projects are excellent.

Review of applications will begin immediately and will continue until the positions are filled. Indication of gender and ethnicity for affirmative action statistical purposes is requested as part of the application. For more information contact **Dr. Mark W. Spong**, Dean of Engineering and Computer Science at [mspong@utdallas.edu](mailto:mspong@utdallas.edu).

Applicants should provide a single PDF file with the following information: (1) resume, (2) statement of research and teaching interests, and (3) full contact information for four, or more, professional references via the **ONLINE APPLICATION FORM** available at <http://go.utdallas.edu/pcu141029>

*The University of Texas at Dallas is an Equal Opportunity/Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability, pregnancy, age, veteran status, genetic information or sexual orientation.*



Massachusetts  
Institute of  
Technology

Come work with us!

## Tenure Track Faculty

**The Department of Brain & Cognitive Sciences (BCS)** (<http://bcs.mit.edu>) and **The Picower Institute for Learning & Memory at MIT** (<http://picower.mit.edu/>) are looking to hire up to three (3) tenure-track faculty at the assistant professor level who work in one or more of the following three (3) areas:

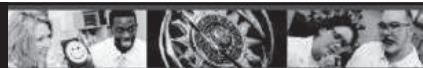
- i) *Computational* approaches to intelligence, cognition or neuroscience; an experimental component to the candidate's research would be viewed as a positive but is not necessary. An affiliation with Electrical Engineering and Computer Science, the Computer Science and Artificial Intelligence Laboratory (CSAIL), or other allied departments is possible.
- ii) *Molecular & cellular*: The Picower Institute is searching for a candidate studying development, function or plasticity of neuronal circuits at the cellular, circuit, and/or systems levels using a multi-faceted approach combining different methodologies and levels of analysis. Candidates with strong cellular/molecular training who are studying development of brain circuits or using stem cell technologies are particularly encouraged to apply.
- iii) *Human cognition and/or cognitive neuroscience* using behavioral methods, especially in the areas of language and/or cognitive development OR using fMRI/neuroscience methods.

Successful applicants are expected to develop and lead independent, internationally competitive research programs and to share in our commitment to excellence in undergraduate and graduate education by teaching courses and mentoring graduate and undergraduate research. PhD must be completed by start day of employment and some postdoctoral training is preferred.

Please submit application materials – cover letter, CV, statement of research and teaching interests and representative reprints – online at <https://academicjobsonline.org/ajob/jobs/4202>. Please state research area in cover letter. To help direct the application, applicants should indicate which of the three areas listed above is their main research area by answering the mandatory questions included in the application. In addition, please arrange to have three letters of recommendation submitted online. Review of applications will begin on November 1, 2014.

MIT is an affirmative action employer, and we encourage applications from women and underrepresented minorities.

<http://web.mit.edu>



## ASSISTANT/ASSOCIATE PROFESSOR

### Department of Molecular Medicine - College of Medicine

We are seeking candidates for a tenure-track position at the Assistant or Associate Professor rank who wish to develop an innovative research program within the broad area of molecular microbiology. The successful candidate will complement or enhance current areas of strength in the department including antimicrobial drug development and target identification, bacterial pathogenesis, molecular virology, protein structure and function, and host response to infection. In addition, qualified candidates studying the role of the microbiome in human disease are encouraged to apply. The Department of Molecular Medicine is located on the energetic and expanding USF Health campus that includes the Colleges of Medicine, Nursing, Pharmacy, and Public Health, the H. Lee Moffitt NCI-designated Comprehensive Cancer Center and the Byrd Alzheimer's Institute. Opportunities for translational research collaboration exist with clinical faculty in the Divisions of Infectious Diseases in the Departments of Internal Medicine and Pediatrics at the USF Morsani College of Medicine as well as the Center for Infection Research in Cancer at the Moffitt Cancer Center. USF is ranked 34rd in total research expenditures and 27th in federal research expenditures for public universities by the National Science Foundation for 2013 with extramural grants and contract activity in excess of \$412 million annually.

Successful candidates are expected to develop and maintain a competitively funded research program, participate in interdisciplinary research addressing molecular microbiology, and educate medical and graduate students. The Department offers institutional salary support commensurate with experience, competitive start-up packages, laboratory space, excellent core facilities, and a collegial atmosphere to promote success. Minimum requirements for an Assistant Professor position include a MD, PhD, or MD/PhD with at least two years postdoctoral experience and current grant funding support. Appointment at the Associate Professor level requires a minimum of five years of experience as an Assistant Professor or equivalent, a nationally-recognized and federally funded research program and experience mentoring graduate students.

All applications must be submitted on line by going to: USF Careers (<http://www.usf.edu/about-usf/work-at-usf.aspx>) and uploading a PDF that includes a cover letter, curriculum vitae, research plan, and statement of teaching philosophy. Separately, request three letters of reference to be sent by mail or e-mail to Dana Pettaway at [dpettawa@health.usf.edu](mailto:dpettawa@health.usf.edu). Review of applicants will begin immediately and will continue until the position is filled.

Microbiology Search Committee, Department of Molecular Medicine  
University of South Florida-Health, 12901 Bruce B. Downs Blvd. MDC7  
Tampa, Florida 33612-4799

[http://health.usf.edu/nocms/research/molecular\\_medicine/](http://health.usf.edu/nocms/research/molecular_medicine/)

The University of South Florida is an Equal Opportunity / Affirmative Action / Equal Access Institution.

For disability accommodations, contact Dana Pettaway at (813) 974-8320 - a minimum of 5 working days in advance. According to Florida law, search records, including applications and search committee meetings, are open to the public.



• TAMPA • ST. PETERSBURG • SARASOTA • MANATEE

Max Planck Institute for Molecular Physiology in Dortmund announces an independent

## Research Group Leader (m/f) position in the area of Synthetic Biology

The Max Planck Institute for Molecular Physiology in Dortmund seeks to appoint a Research Group Leader in the area of Synthetic Biology. The ideal candidate should have a research program that is centred on the reconstruction of self-organising, biochemical systems that mimic morphogenesis of cells. Specifically projects that centre on membrane-protein interaction are favoured. Applicants are expected to establish an internationally recognized independent research program and are encouraged to collaborate with the groups contributing to the MaxSynBio network, with other groups and Departments at the MPI Dortmund, the TU Dortmund University and the Chemical Genomics Centre in Dortmund.

The candidate should hold an advanced degree in either physics, (bio)chemistry, biology or a related discipline and have at least two years post-doctoral experience. The contract is for a Research Group Leader position at the Assistant Professor level and limited to 5 years with the possibility of extension after successful evaluation and considering labour law-related and financial requirements. The support includes average salary group up to E15 or W2, where appropriate, based on the German Federal Salary Scale, as well as support for an independent group (consumables and staff).

The Max Planck Institute for Molecular Physiology is committed to research on how complex, systemic properties at a higher scale emerge from the dynamics of local interactions of molecular components in living systems. The four existing Departments develop and use various approaches to address fundamental questions arising on the different scales of biological systems – from molecules to cells and tissues.

The Max Planck Society is committed to increasing the number of individuals with disabilities in its workforce and therefore encourages applications from such qualified individuals.

The Max Planck Society seeks to increase the number of women in those areas where they are underrepresented and therefore explicitly encourages women to apply.

Applications, including curriculum vitae, list of publications, and an outline of previous (1 page) and future research objectives (2 pages) as well as names of referees should be sent electronically as one file by **December 15<sup>th</sup>, 2014** to [synthbiol@mpi-dortmund.mpg.de](mailto:synthbiol@mpi-dortmund.mpg.de).

Short-listed candidates will be invited to a one-day symposium, at which they will have the opportunity to present their research. Date and place of the one-day symposium: **February 5<sup>th</sup>, 2015**, Max Planck Institute for Molecular Physiology, Dortmund, Germany. For further information please contact: [synthbiol@mpi-dortmund.mpg.de](mailto:synthbiol@mpi-dortmund.mpg.de)



MAX-PLANCK-GESELLSCHAFT



## Assistant/Associate Professor Laboratory Investigator: Ovarian Cancer

The Department of Medical Oncology at the Dana-Farber Cancer Institute (DFCI), the Gynecologic Oncology Program of the Susan F. Smith Center for Women's Cancers, and the Brigham and Women's Hospital, invite applications for a full-time appointment at the Assistant or Associate Professor level. This individual will develop an independent laboratory-based translational research program focused on ovarian cancer. The research program will interface directly with the translational and clinical research efforts within the Gynecologic Oncology program at DFCI as well as other laboratories at DFCI. Candidates with interests in the genomic basis of ovarian cancer, ovarian cancer biology, and/or immunology as well as research engaged in pre-clinical development of new therapeutic approaches are especially encouraged to apply. The candidate must have an MD and/or PhD and a proven track record of outstanding laboratory research.

The candidate will work principally at the Dana-Farber Cancer Institute and the Brigham and Women's Hospital. Appointment as Assistant or Associate Professor at the Harvard Medical School will be commensurate with experience, training and achievements. Salary and benefits will be competitive with other institutions. Candidates should be board certified or board eligible in Medical Oncology and Internal Medicine. Dana-Farber Cancer Institute is an NCI-designated Comprehensive Cancer Center.

**Interested candidates must submit a curriculum vitae, a research plan and 3 letters of reference to: Ursula Matulonis, M.D., Director, Gynecologic Oncology Program, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215. Please send submissions via email to: [umatulonis@partners.org](mailto:umatulonis@partners.org).**



HARVARD  
MEDICAL SCHOOL



BRIGHAM AND  
WOMEN'S HOSPITAL

We are an equal opportunity employer and all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability status, protected veteran status, or any other characteristic protected by law.

## Assistant Professor

Departments of Cancer Biology  
Dana-Farber Cancer Institute and  
Cell Biology  
Harvard Medical School

The Departments of Cancer Biology at the Dana-Farber Cancer Institute and Cell Biology at Harvard Medical School invite applicants for a tenure-track faculty position at the rank of Assistant Professor or Associate Professor. We are seeking individuals with a demonstrated potential for imaginative research and who propose to work on exciting problems in any area of Molecular Metabolism. We are especially interested in candidates who utilize metabolomics, proteomics or both to uncover novel metabolic pathways in health and diseased states. The successful candidate will be expected to direct innovative and independent research and participate in the teaching of graduate and/or medical students. Our highly interactive environment provides the opportunity to engage and collaborate with other dedicated researchers both within the Division of Metabolism and Chronic Disease of the Cancer Biology Department (chaired by Bruce Spiegelman) and throughout the diverse Harvard research community. Significant scholarly and scientific resources will be made available for this appointment. Applicants will be housed in new space at the Dana-Farber Cancer Institute. For further information about our Department, please see our web page: <http://www.dana-farber.org/Research/Departments-and-Centers/Department-of-Cancer-Biology.aspx>

**Applicants should submit electronic copies of their curriculum vitae, a description of research accomplishments and future research interests (three pages maximum), and ask at least three references to provide letters of recommendation. These materials should be submitted using the following link: <https://academicpositions.harvard.edu/postings/5815>. Please contact Kim Wilkinson ([Kim\\_Wilkinson@dfci.harvard.edu](mailto:Kim_Wilkinson@dfci.harvard.edu)) with any questions regarding submission of documents.**

**Applications must be received by: February 28, 2015.**

We are an equal opportunity employer and all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability status, protected veteran status, or any other characteristic protected by law.



DANA-FARBER  
CANCER INSTITUTE



HARVARD  
MEDICAL SCHOOL

## Florida State University

### Strategic Faculty Recruitment in Energy & Materials: Materials Characterization and Energy Devices



Florida State University is continuing its interdisciplinary faculty hiring initiative in **Energy & Materials**. During the 2014-15 academic year the University seeks to *recruit several established, mid-career faculty members with the equivalent experience and accomplishments of Associate Professors*. We invite applications from researchers who develop and utilize advanced methods for **fundamental materials characterization** as well as those who work in the area of **design, prototyping and fabrication of novel energy devices**. Academic department of appointment is open. These are nine month, tenured/tenure-track faculty positions. Successful candidates are expected to have a synergistic impact on existing research programs in the University's departments and interdisciplinary centers, to contribute to scientific leadership in the growing materials community and to participate in teaching and mentoring at the undergraduate and graduate levels. Successful candidates will be offered a highly competitive salary and start-up package, high quality research space and access to state-of-the-art instrumentation, computing and facilities in academic and interdisciplinary units. Related strengths at Florida State University include departments in the College of Arts & Sciences and the College of Engineering. Complementing these programs are interactive centers including the National High Magnetic Field Laboratory, the Applied Superconductivity Center, the High Performance Materials Institute, the Aero-Propulsion, Mechatronics & Energy Center, and the Center for Advanced Power Systems. Linking these colleges and centers is a new Ph.D. program in Materials Science & Engineering (<http://www.materials.fsu.edu/>) complementing robust department-based doctoral programs in materials and related areas.

Applicants are asked to provide a single document in PDF format containing a letter of application, a curriculum vitae, a two page narrative describing their research interests and plans, and a brief teaching statement. Applications must be sent electronically to [materials2014.search@fsu.edu](mailto:materials2014.search@fsu.edu). Highly ranked applicants will be contacted at a later date for the purpose of requesting letters of recommendation. Review of applications will begin **January 2, 2015** and will continue until the positions are filled.

*Florida State University is committed to the diversity of its faculty, staff, and students, and to sustaining a work and learning environment that is inclusive. Women, minorities, and people with disabilities are encouraged to apply. FSU is an Equal Opportunity/Access/Affirmative Action Employer.*

## Northeastern University

### College of Engineering

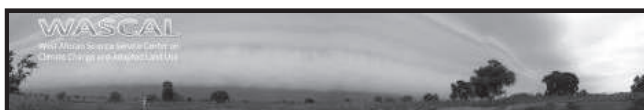
With **44** tenure-track hires since 2011, **8** federally-funded research centers, and a bioengineering department established in **2013**, Northeastern's College of Engineering is in a period of dynamic growth. Our emphasis on interdisciplinary, use-inspired research—tied to Northeastern's unique history of industry collaboration via the university's signature cooperative education program—enables partnerships with academic institutions, medical research centers, and companies near our centrally located Boston campus and around the globe.

**The college seeks outstanding faculty candidates, with emphasis on bio-related, interdisciplinary expertise.**

Successful applicants will lead internationally recognized research programs aligned with one or more of the college's strategic research initiatives. Particular consideration will be given to candidates at the associate or full professor level; exceptional candidates at the assistant professor level will also be considered.

**Learn more and apply at  
[coe.neu.edu/faculty/positions](http://coe.neu.edu/faculty/positions)**

Northeastern is an Affirmative Action/Equal Opportunity educator and employer committed to excellence through diversity.



**WASCAL**  
West African Science Service Center on  
Climate Change and Adapted Land Use

### Senior Data Manager

**WASCAL, the West African Science Service Center on Climate Change and Adapted Land Use**, is a research and education-focused inter-governmental institute designed to help tackle the climate change challenge and thereby enhance the resilience of human environmental systems. WASCAL operates a Competence Center (CC) in Ouagadougou, Burkina Faso, for its climate research activities. WASCAL is seeking to fill the position of a Senior Data Manager to be based at its CC.

The successful applicant will supervise the operation of the data infrastructure at the CC, which manages heterogeneous research data sets, and co-ordinate its further development. He/she will act as the primary contact person for the scientific staff with regard to all data management issues at the CC. Major responsibilities of the position are the provision of senior professional leadership to IT systems, including data management software, and the development of concepts and data management plans for system extensions in terms of software architecture and user interface design according to changing demands of research carried out in the framework of WASCAL.

The applicant should have a PhD in (geo-) informatics/computer science, or in another field closely related to management of research data, together with sound experiences (minimum 5 years) with data management systems operating on distributed data infrastructures. He/she has good knowledge on database modeling and management as well as on the concepts and standards used in Spatial Data Infrastructures (OGC web-services). Excellent communication and team leadership skills with excellent English and basic French speaking and writing are expected. This is an international staff position based in Ouagadougou, Burkina Faso, and requires to travel within the WASCAL region and to Germany.

Detailed job description and application procedure at: <https://icg4wascal.icg.kfa-juelich.de/downloads/vacancy-senior-data-manager-2014/view> (download the PDF)

## GROUP LEADER POSITION

### in plant non-coding RNA biology Strasbourg, France

A position for a group leader in the field of non-coding RNA biology in plants is opened at IBMP (Institut de Biologie Moléculaire des Plantes), Strasbourg, France. IBMP is a leading plant biology research institute and provides a creative, stimulating and international environment with staffed core facilities (<http://ibmp.u-strasbg.fr>).

The group leader position is opened in the frame of the NetRNA Laboratory of Excellence (LabEx) (<http://labex-ibmc.u-strasbg.fr/NetRNA/>), a consortium of eleven teams from two nearby institutes (IBMP and IBMC). The general objective of the NetRNA consortium is to advance knowledge on functions of regulatory RNAs, including noncoding RNAs (ncRNAs), and their mechanism of action in a variety of pathogens (bacteria, viruses, parasites) and their hosts, including plants, insects, and mammals.

Current RNA-based research at the IBMP focuses on RNA Quality control (RQC) and the mechanisms and roles of small RNA-mediated silencing, as well as post-translational regulations of the silencing machineries. Applications should ideally cover a new topic and/or make use of alternative model systems than *Arabidopsis thaliana*, including single cell organisms (e.g. *Chlamydomonas*).

The offer includes lab space for 6-8 persons, access to common equipment and on-site facilities and core funding of 550k€ over a three year period covering salaries and consumables. Candidates should have an internationally competitive publication record. Applicants should send a CV with a list of publications, a 3-4 page synopsis summarizing accomplishments and future research plans, names and contacts of at least three academic references to [ibmp-callnetrna@unistra.fr](mailto:ibmp-callnetrna@unistra.fr).



# Professor (all levels) in Metabolic Engineering

King Abdullah University of Science and Technology (KAUST) (<http://www.kaust.edu.sa>) is seeking a highly motivated and skilled faculty member whose research focuses on metabolic pathway design, optimization, simulation, and engineering aimed at industrial biotechnology applications.

KAUST is an international, graduate-level research university dedicated to advancing science and technology through interdisciplinary research, education, and innovation. Located on the shores of the Red Sea in Saudi Arabia, KAUST offers superb research facilities, generous assured research funding, and internationally competitive salaries, attracting top international faculty, scientists, engineers, and students to conduct fundamental and goal-oriented research to address the world's pressing scientific and technological challenges in the areas of food, water, energy, and the environment.

The successful applicant is expected to develop strong research in computational design and optimization of metabolic pathways. This research will be complemented by extensive experimental research for the validation of computational designs. The faculty member is also expected to establish an experimental laboratory. The faculty member will be part of the Computational Bioscience Research Center (CBRC) within the Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division. The position will remain open until filled.

## Requirements:

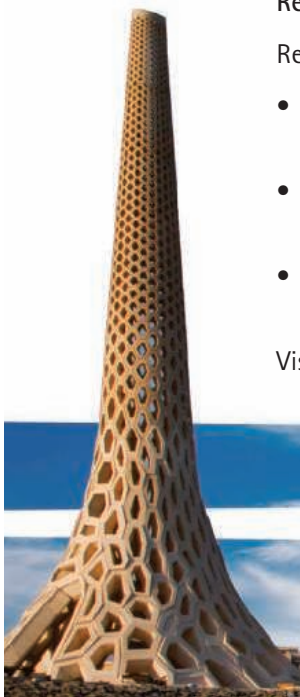
PhD or equivalent in a relevant discipline. Candidates should be well-established within the research field. They should demonstrate original research and experience at the highest international level.

## Responsibilities and tasks:

Research competence in the following areas is preferred:

- Computational design, optimization, and simulation of metabolic pathways for the bio-production of important chemicals
- Bioinformatics analysis of transcriptomic and metabolomic data for improvement of host organisms used for the production of biochemicals
- Novel high throughput computational or experimental tools for the optimization of metabolic pathways for production of biochemicals at high yield.

Visit <http://apptrkr.com/536231> to apply.





Johns Hopkins Bloomberg School of Public Health

### Vector Biologist

The Malaria Research Institute and the Department of Molecular Microbiology and Immunology in the Johns Hopkins Bloomberg School of Public Health invite applications for a tenure track faculty position at the **Assistant Professor** level in **Vector Biology**.

Extensive insectary and core facilities are available to successful applicants, including the JHMRI field site in Zambia. Interest in field investigation is encouraged, but not required. A competitive start-up package, salary and benefits will be provided. Individuals will be expected to develop independent research programs within an interactive environment of investigators interested in the pathogenesis of microbial diseases. There will be opportunities for teaching and training graduate students. Applicants must have a PhD, MD, or equivalent degree and appropriate post-doctoral experience.

Applicants should provide curriculum vitae, description of research interests and the names of at least three references by **January 15, 2015**.

For more information, or to submit an application, please contact Genevieve Williams, Program Manager, Johns Hopkins Malaria Research Institute, Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, E5141, 615 North Wolfe Street, Baltimore, MD, 21205. Voice 410-614-4883, fax 410-502-5884, Genevieve.Williams@jhu.edu.

Established in 2001, the JHMRI supports a broad program of basic research targeting novel approaches to malaria treatment and control. For information on JHMRI faculty, research programs, fellowships, core resources, and the Malaria Institute at Macha, please visit our website, <http://malaria.jhsph.edu>

*The Johns Hopkins University actively encourages interest from women and minorities and is an Affirmative Action/Equal Opportunity Employer.*



JOHNS HOPKINS  
BLOOMBERG SCHOOL  
of PUBLIC HEALTH



## AAAS is here – promoting universal science literacy.

In 1985, AAAS founded Project 2061 with the goal of helping all Americans become literate in science, mathematics, and technology. With its landmark publications *Science for All Americans* and *Benchmarks for Science Literacy*, Project 2061 set out recommendations for what all students should know and be able to do in science, mathematics, and technology by the time they graduate from high school.

As a AAAS member, your dues help support Project 2061 as it works to improve science education. If you are not yet a member, join us. Together we can make a difference.

To learn more, visit [aaas.org/plusyou/project2061](http://aaas.org/plusyou/project2061)



### Head, Department of Biochemistry

We are looking for a dynamic individual with a creative and inspiring vision for biochemistry research and education. Our forward-thinking, diverse, and internationally recognized department has 20 full-time faculty members with active research programs funded by the NIH, NSF, DOE, and ACS, as well as industry. These research programs encompass a wide variety of topics including metabolic and natural product biochemistry, cancer biochemistry, epigenetics and gene expression, proteomics and metabolomics, structural biology and macromolecular complexes. The majority of our faculty are members of the NCI designated Purdue University Center for Cancer Research and/or play key roles in Purdue's Plant Sciences and Biofuels initiatives. The Department is dedicated to providing an exceptional undergraduate education that emphasizes experiential learning. Excellence in graduate education is a central goal of the department and is evidenced by a well-established graduate program; many biochemistry faculty also participate in the University-wide interdisciplinary life science graduate program (PULSe). Purdue University is committed to continued promotion of excellence in biochemistry research, and educational and scientific innovation.

See <http://www.biochem.purdue.edu> for full position description and application procedures.

*Purdue University is an EEO/AA Employer fully committed to achieving a diverse workforce. All individuals, including minorities, women, individuals with disabilities, and protected veterans are encouraged to apply.*



## AAAS is here – helping scientists achieve career success.

Every month, over 400,000 students and scientists visit ScienceCareers.org in search of the information, advice, and opportunities they need to take the next step in their careers.

A complete career resource, free to the public, *Science Careers* offers a suite of tools and services developed specifically for scientists. With hundreds of career development articles, webinars and downloadable booklets filled with practical advice, a community forum providing answers to career questions, and thousands of job listings in academia, government, and industry, *Science Careers* has helped countless individuals prepare themselves for successful careers.

As a AAAS member, your dues help AAAS make this service freely available to the scientific community. If you're not a member, join us. Together we can make a difference.

To learn more, visit  
[aaas.org/plusyou/sciencecareers](http://aaas.org/plusyou/sciencecareers)







## Qingdao Technological University Recruits for High-end Talents

Qingdao Technological University is a multiversity includes Science, Engineering, Economics, Management, Literary, Law and Arts. Particularly at Science and Engineering, Civil Construction, Machine Manufacturing, and Environmental Energy.

Enjoys a 60 years history, the university builds a higher education system with undergraduates, masters and doctors. The university now has three campuses: Shibei, Huangdao, and Feixian, which accommodates 19 academic schools, 59 undergraduates programs, 18 level-1 disciplines offering master's degree programs, 1 first-level discipline offering doctoral degree programs, 7 second-level disciplines offering doctoral degree programs, 2 first-level discipline post-doc research stations, and 25 provincial key disciplines, key labs and engineering technical research centers. Now the University has 2215 faculty and staff members, and over thirty thousand students.

In accordance with the university development, now we are recruiting high-end talents globally.

### Disciplines

Civil Engineering, Mechanical Engineering, Environmental Science and Engineering, Architecture, Rural and Urban Planning, Landscape Architecture, Design Planning, Communication and Transportation Engineering, Safety Science and Engineering, Computer Science and Technology, Software Engineering, Information and Communication Engineering, Control Science and Engineering, Materials Science and Engineering, Mechanics, Management science, Business Administration, and Applied Economics.

### Requirements

Applicants should have doctorate, also hold the position of assistant professor, associate professor, and professor at leading universities or research institutions overseas. Professionals and talented management at well-known enterprise or financing institution would be acceptable as well.

### Contact information

For more information, please check on our main site: <http://www.qtech.edu.cn>

Tel: +86-0532-85071322

E-mail: [qdlgrc@163.com](mailto:qdlgrc@163.com)



西南交通大学  
Southwest Jiaotong University

Southwest Jiaotong University, P.R.China  
Anticipates Your Working Application

Southwest Jiaotong University (SWJTU), founded in 1896, situates itself in Chengdu, the provincial capital of Sichuan. It is a national key multidisciplinary "211" and "985 Feature" Projects university directly under the jurisdiction of the Ministry of Education, featuring engineering and a comprehensive range of study programs and research disciplines spreading across more than 20 faculties and institutes/centers. Boasting a complete Bachelor-Master-Doctor education system with more than 2,500 members of academic staff, our school also owns 2 first-level national key disciplines, 2 supplementary first-level national key disciplines (in their establishment), 15 first-level doctoral programs, 43 first-level master programs, 75 key undergraduate programs, 10 post-doctoral stations and more than 40 key laboratories at national and provincial levels.

Our university is currently implementing the strategy of "developing and strengthening the university by introducing and cultivating talents". Therefore, we sincerely look forward to your working application.

More information available at <http://www.swjtu.edu.cn/>

### I. Positions and Requirements

#### A. High-level Leading Talents

It is required that candidates be listed in national top talents programs such as *Program of Global Experts*, *Top Talents of National Special Support Program*, *"Chang Jiang Scholars"*, *China National Funds for Distinguished Young Scientists* and *National Award for Distinguished Teacher*.

Candidates are supposed to be no more than 50 years old. The limitation could be extended in the most-needed areas of disciplinary development.

Candidates who work in high-level universities/institutes and reach the above requirements are supposed to be no more than 45 years old.

#### B. Young Leading Scholars

Candidates are supposed to be listed in or qualified to apply for the following programs:

• *National Thousand Young Talents Program*

• *The Top Young Talents of National Special Support Program (Program for Supporting Top Young Talents)*

• *Science Foundation for the Excellent Youth Scholars*

Candidates should have good team spirit and leadership, outstanding academic achievements, broad academic vision and international cooperation experience and have the potential of being a leading academic researcher.

#### C. Excellent Young Academic Backbones

Candidates under 40 years old are expected to graduate from high-level universities/institutes either in China or other countries. Those who are professors, associate professors and other equal talents from high-level universities/institutes overseas could be employed as professors and associate professors as well.

#### D. Excellent Doctors and Post Doctoral Fellows

Candidates under 35 years old are supposed to be excellent academic researchers from high-level universities either in China or other countries.

### II. Treatments

The candidates will be provided with competitive salaries and welfares that include settling-in allowance, subsidy of rental residence, start-up funds of scientific research, assistance in establishing scientific platform and research group as well as international-level training and promotion. As for outstanding returnees, we can offer further or specific treatments that can be discussed personally.

### III. Contact us:

Contacts: Ye ZENG & Yinchuan LI

Telephone number: 86-28-66366202

Email: [talent@swjtu.edu.cn](mailto:talent@swjtu.edu.cn)

Address: Human Resources Department of SWJTU, the western park of high-tech zone, Chengdu, Sichuan, P.R.China, 611756

<http://www.swjtu.edu.cn/>

## Job Vacancies in China's Universities



赛尔互联  
CERNET

Science Careers  
From the Journal Science

赛尔互联:《科学》在中国大陆高校人才引进领域独家合作伙伴  
CERNET is Science's exclusive agent for recruitment advertisement service in mainland China universities and colleges

## China's Rapid Development — More Opportunities

### Xi'an Jiaotong University (Xi'an, China)



Faculty positions in Department of Sustainable Energy:  
Mission: To pioneer forward-looking education and undertake cutting-edge research in energy science and technologies to support sustainable development.

Faculty Positions in Department of Sustainable Materials:  
Mission: To provide a broad-based education in materials science and engineering for sustainability and undertake cutting-edge research in the field to support sustainable development.

Faculty Positions in Department of Sustainable Systems:  
Mission: To provide an education and research platform to study interrelationships among eco-nomic and urban development, governance, public policy and regional, national and global ecosystem, focusing on the social and physical systems needed for sustainability.

For further information about JSSD, please visit the following websites:  
<http://jssd.xjtu.edu.cn> or <http://jssd.usf.hk>

### Beijing Institute of Technology (Beijing, China)

Open Senior Faculty Positions offered: Equipment Science & Technology Engineering/ Mechanical Engineering/ Information and Communication Engineering/ Mathematics...

### Peking University (Beijing, China)

The School of Life Sciences (SLS) at Peking University invites applications for multiple faculty positions at tenure-track assistant professor, associate professor and full professor levels.

For more details,  
visit <http://www.acabridge.cn/>

广告联系: 赵佳 [zhaojia@cernet.com](mailto:zhaojia@cernet.com)  
+86 10 62603373

中国教育网 是赛尔互联旗下品牌



暨南大学  
JINAN UNIVERSITY

## Faculty Position Available in Jinan University, Guangzhou, South China



### Faculty Positions Available in School of Medicine, Jinan University

School of Medicine, Jinan University, is to become a world-class medical school. Now we are searching Senior Scientists (full professor or associate professor) for the Department of Anatomy, the Department of Physiology, the Department of Pathology, the Department of Public health, the Department of Nursing, the Department of Dentistry and so on.

#### Applicants should:

- \* have a doctoral degree in the relevant research field, and a proven record of research excellence;
- \* have team work spirits and coordination skills, leadership ability;
- \* top-priority will be given to candidates hold a professorship/associate professorship in a renowned overseas university (or research institutes).

#### Compensation and benefit:

- \* general compensation and benefits are on line with current standards of Jinan University
- \* generous start-up packages
- \* successful applicants who meet relevant criteria will be nominated for the various talent programs of the Chinese government and the Guangdong Province. Please submit your curriculum vitae including a cover letter declaring the targeted research area via post or email to:

Ms. Emily Huang  
(E-mail: [thyzh@email.jnu.edu.cn](mailto:thyzh@email.jnu.edu.cn); Tel: +86-20-85228912)  
Mr. Huanmin Luo  
(E-mail: [tlhm@jnu.edu.cn](mailto:tlhm@jnu.edu.cn); Tel: +86-20-85220500)



### Faculty Positions Available in Department of Optoelectronic Engineering, Jinan University

University fellowship is available. We seek highly motivated young researchers with solid strength in the photonics research areas, especially in ultrafast laser optics or novel optical fiber lasers.

Fellowship candidates should possess an earned Ph.D degree in physical science, or optical science or engineering. Candidates under 35-years old and with degree from world-famous universities are preferred. Excellent English writing and oral communication skills are required. The fellowship positions are in the following areas:

#### I Ultrafast laser optics:

An essential research objective in our group is focused on making high power femtosecond laser with novel laser crystals and laser design. We seek researchers with a strong experimental expertise and theoretical knowledge in the ultrafast laser optics, in particular femtosecond laser design and nonlinear optics.

#### II Optical fiber lasers:

We aim at making optical fiber laser that can produce either high power femtosecond pulses, as well as short cavity fiber lasers using novel specialty optical fibers. We seek candidates with strong experience in fiber optics and laser physics.

Applicants should send a full CV via email to:

Mr. Zhenqiang Chen  
(Email: [tzqchen@jnu.edu.cn](mailto:tzqchen@jnu.edu.cn);  
Tel: +86-20-85220484 ext 417)





**FACULTY POSITION IN MICROBIOLOGY**  
The University of Oklahoma  
Department of Microbiology and Plant Biology

The Department of Microbiology and Plant Biology at the University of Oklahoma invites applications for a tenure-track faculty position at the **ASSISTANT PROFESSOR** level beginning August 2015. A candidate is sought who can bring synergistic strengths in the areas of systems biology, synthetic biology, microbial interactions, microbial physiology, or bioinformatics toward addressing important questions in microbiology. The department has core strengths in the areas of environmental microbiology, biodegradation, and energy, especially as they relate to anaerobic systems. Institutional priorities include research related to water, biotechnology, energy, climate and cancer biology, and the successful candidate will be encouraged to participate in collaborative, cross-disciplinary, and cross-departmental research and initiatives. Candidates must have a Ph.D. in Microbiology or a related discipline, a strong record of research publication, post-doctoral experience, and the ability to develop and maintain an independent, extramurally funded research program. The successful candidate will contribute to departmental teaching and training of undergraduate and graduate students in support of our curricula. Applications should be submitted as a single PDF-format file to e-mail: [mpbiofacultysearch@ou.edu](mailto:mpbiofacultysearch@ou.edu) and should include a current curriculum vitae, up to five representative reprints, and a statement of research plans as well as a statement of teaching interests and philosophy. Three to five letters of reference should be sent directly to the e-mail address above. Review of applications will begin January 1, 2015, and will continue until the position is filled. *The University of Oklahoma is an Affirmative Action/Equal Opportunity Employer and encourages diversity in the workplace. Protected veterans and individuals with disabilities are encouraged to apply.*

**FACULTY POSITION**  
Johns Hopkins University  
School of Medicine  
Institute for Cell Engineering

The Institute for Cell Engineering invites applications from outstanding individuals with creative, rigorous, and integrative research approaches to enhance its cell engineering investigational strengths in immunology, stem cell biology, neurosciences, and vascular biology. For additional information about the institute, visit website: [http://www.hopkinsmedicine.org/institute\\_cell\\_engineering/](http://www.hopkinsmedicine.org/institute_cell_engineering/). Candidates should have an M.D. and/or a Ph.D. degree with appropriate post-doctoral experience and an outstanding publication record. Primary department affiliation will be determined by the applicant's qualifications and by relevance of the applicant's research program to departmental initiatives. The successful candidate will have experience in any aspect of stem cell biology or related field. Special attention will be given to investigators in the areas of vascular biology or organogenesis.

To apply, submit curriculum vitae, three letters of reference, copies of relevant publications and a concise description of research interests and research plans (up to three pages) to e-mail: [icsearch@jhmi.edu](mailto:icsearch@jhmi.edu) to the attention of **Ted M. Dawson**, Director, Institute for Cell Engineering. Applications will be assessed on an ongoing basis and the deadline for submission is December 15, 2014. The appointment is expected to be made in 2015.

*The John Hopkins University School of Medicine is an Affirmative Action/Equal Opportunity Employer that embraces diversity.*

Get your questions answered.  
Careers Forum  
**ScienceCareers.org**

# Download your free copy today.

[ScienceCareers.org/booklets](http://ScienceCareers.org/booklets)



From technology specialists to patent attorneys to policy advisers, learn more about the types of careers that scientists can pursue and the skills needed in order to succeed in nonresearch careers.

**ScienceCareers**  
FROM THE JOURNAL SCIENCE AAAS



## Learn more and conduct your job search the easy way.

- Search thousands of job postings
- Create job alerts based on your criteria
- Get career advice from our Career Forum experts
- Download career advice articles and webinars
- Complete an individual development plan at "myIDP"

Target your job search using relevant resources on **ScienceCareers.org**.

**ScienceCareers**  
FROM THE JOURNAL SCIENCE 



Boston, MA

**Northeastern**

## Assistant/Associate Professor - Ecological & Evolutionary Genomics

*Northeastern University,  
Marine Science Center*

**REQUISITION NUMBER:** FTFR000549

**DIVISION/COLLEGE:** College of Science

**FT/PT:** Full Time

**POSITION SUMMARY:** The Department of Marine and Environmental Sciences at Northeastern University, Boston, Massachusetts, invites applications for a tenure-track appointment in Ecological and Evolutionary Genomics. We seek candidates working with non-model organisms (marine or terrestrial) and using next-generation sequencing techniques to address fundamental questions related to environmental change, particularly in coastal ecosystems. The new hire's research laboratory will be based at the Marine Science Center in Nahant, Massachusetts and the successful candidate will have the opportunity to make important contributions to the Marine Science Center's new partnership with the Ocean Genome Legacy.

The anticipated start date will be the fall semester of 2015. A competitive start-up package will be provided.

The successful candidate will be expected to teach at both the undergraduate and graduate level, conduct an independent, externally funded research program, and make strong contributions to the College of Science's Urban Coastal Sustainability Initiative. The department has 17 full-time faculty members and administers or contributes to undergraduate programs in Environmental Science and Studies, Marine Biology, and Biology. It hosts a Ph.D. program in Ecology, Evolution and Marine Biology, and Masters programs in Marine Biology and Bioinformatics.

**QUALIFICATIONS:** Applicants must have a PhD in Marine Science, Ecology/Evolutionary Biology or a related field prior to the start of the appointment, a strong publication record and evidence of funding success.

**ADDITIONAL INFORMATION:** Review of applications will begin in the fall of 2014 and will continue until the position is filled. Please contact the Chair of the Search Committee, Professor Geoffrey C. Trussell, at 781-581-7370 (ext. 300) or via email (g.trussell@neu.edu) for questions about the search.

Candidate's should be prepared to attach their CV, Cover Letter, Research Statement, Teaching Statement and 3 Sample publications and contact information for three references.

**How To Apply:** To apply, visit Careers at Northeastern at <http://apptkr.com/542130> You may also apply by visiting the College of Science website at <http://www.northeastern.edu/cos> and clicking the Open Faculty Positions tab.

Northeastern University is an Equal Opportunity, Affirmative Action Educational Institution and Employer, Title IX University.



By Joseph Swift

# The point of it all

It seems like a rite of passage now for Ph.D. students: confronting the harsh prospects of academic careers. I'm just at the beginning of my doctorate at New York University (NYU) in New York City, but already my peers are concerned about how to survive within academic institutions. Our career paths are described as risky undertakings, littered with pitfalls and wrong turns, and the journey can often lead to a dead end. At first I shared their anxiety, but I started to wonder about how this would affect my development as a scientist. How can one produce the best science while focusing on survival? If students remain in a constant state of competitiveness, they may end up trading their creativity for a shortsighted scramble for results. Would merely surviving fulfill my goals?

I decided it was better not to overthink the whole survival thing. Negativity paired with a good imagination can be crippling. As with experiments, sometimes it is best to relax expectations until you see the results.

As I was thinking along these lines, I was offered a rotation project at NYU Abu Dhabi in the United Arab Emirates. The goal was to develop algae strains that could be used as a cheap source of biofuel. As in all rotation projects, I was given a taste of the research, but I didn't have time to really sink my teeth in. I embraced the opportunity because I wanted to work on an environmental biotech project, but I didn't expect to save the world.

My 9 weeks in Abu Dhabi provided something unexpected: a renewed sense of purpose.

Abu Dhabi is a rapidly evolving urban center; think Chinese megacities. What makes Abu Dhabi remarkable is that it survives in the middle of a desert. What little arable land and fresh water Abu Dhabi has cannot provide for a city that size. Resources must instead be created from scratch.

I'm passionate about environmental biotechnology, but up until that point, my exposure to the subject had been limited to the scientific literature and some small adventures in the lab. The Abu Dhabi project helped me understand what environmental biotech strives to achieve. It became clear that innovations in algal wastewater treatment and salt-tolerant crop breeding could be real game changers in an environment that provides so little. Many Western cities also depend on advanced infrastructure to



***“The Abu Dhabi project helped me understand what environmental biotech strives to achieve.”***

stay afloat, but in Abu Dhabi the need was much more obvious.

This forced me to think differently about my sources of motivation. I became less concerned about how I was going to survive or what I was going to do after my Ph.D. I got more interested in the potential impacts of environmental biotechnology. I suppose the effect might be similar to what medical scientists who work at hospitals experience: Proximity to the site of application inspires a certain kind of urgency. This urgency, I believe, is better for science than that provoked by the academic rat race.

How does Ph.D. research affect the world? As obvious as it sounds, this question is often overlooked during grad school. It's easy to lose sight of the larger purpose of your research, especially when you are

distracted by concerns about survival. I suspect this distraction is present at other career stages as well.

Now that I'm back in New York, the experience in Abu Dhabi seems a bit like a desert mirage, but it continues to influence my thinking. I got a glimpse of how my environmental biotechnology research could have a wider impact.

Wrestling with a research question is a rewarding intellectual exercise, but an appreciation of the broader context of the research can enrich it. These multiple sources of motivation leave little room for career-related anxiety. ■

*Joseph Swift is a Ph.D. student studying plant systems biology at NYU. For more on life and careers, visit [www.sciencecareers.org](http://www.sciencecareers.org). Got an interesting career story? Send it to [SciCareerEditor@aaas.org](mailto:SciCareerEditor@aaas.org).*